



The State of the Gate: A Description of Instructional Practice in Algebra in Five Urban Districts

Citation

Litke, Erica G. 2015. The State of the Gate: A Description of Instructional Practice in Algebra in Five Urban Districts. Doctoral dissertation, Harvard Graduate School of Education.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:16461050>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The State of the Gate:
A Description of Instructional Practice in Algebra in Five Urban Districts

Erica G. Litke

Dr. Heather C. Hill
Dr. Jon R. Star
Dr. David K. Cohen

A Thesis Presented to the Faculty
of the Graduate School of Education of Harvard University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

2015

©2015
Erica G. Litke
All Rights Reserved

To Nathan, Oscar, and Elliot. You are my heart.

Acknowledgments

This dissertation would not be possible without the support, guidance, and encouragement of my committee, all of whose influence is readily apparent in the final product. My advisor, Heather Hill, has been a true mentor in every sense of the word and my work is immeasurably better for her care, attention, keen editing, and tough love. She has taught me to look deeply and clearly at instructional practice while still trusting in my own expertise and experience. She brings to her work a respect for the craft of teaching that is inspirational. From day one, she provided guidance and countenance, nudging me along but allowing me to find my own way.

I am grateful for Jon Star's generosity, insightful feedback, and depth of knowledge of mathematics education. Our interactions have deepened my thinking about the role and promise of the teaching of procedures. His thoughtful advice at crucial junctures has been indispensable.

David Cohen's masterful, well-placed, and well-timed questions moved my work forward. He has taught me to understand contemporary instructional practice in the context of the history and organization of schooling. I am deeply appreciate of his support, advice, and belief in me.

Additionally, I benefitted from the generosity of a number of colleagues. Thank you to Johanna Barmore, Samantha Rabinowicz, and Delaina Martin who worked with me for many months to develop the QIPA. There would be no dissertation without you. Leah Ziegler enthusiastically jumped in and helped score lessons. My colleagues at NCTE, especially Eric Anderson, David Blazar, and Mark Chin, provided technical and

analytic assistance. Thank you to ICPSR at the University of Michigan, especially Johanna Bleckman, for facilitating access to the data.

The seeds of this dissertation were planted at East Side Community High School. Thank you to the East Side staff, especially Mark Federman, Tom Mullen, and Mohammed Aminyar, for existence proof of a culture of strong mathematics instruction. To my students, especially Pebbles, Ashley, Britni, Jillian, Rosalba, Jennifer, Shannon, Nadine, and Courtnie: thank you for inspiring this work.

I was fortunate to be supported through my time at HGSE by an incredible array of mentors. Thank you to John Willett, Susan Moore Johnson, Dick Murnane, John Diamond, and Jal Mehta, each of whom took a vested interest in my success and supported my intellectual and professional growth.

My doctoral experience was deeply enriched by my amazing colleagues and friends, whose encouragement, insight, and humor got me to the finish line. I am especially grateful for my writing group, Bill Johnston and Nikhit D'Sa, who are incredible scholars and dear friends. Thank you to Doc09, especially Adrienne Keene, Monica Ng, Adrienne Mundy-Shepard, Steve Hoffman, John McLaughlin, and Maya Weilundemo, and to my *Harvard Educational Review* family, especially Janine de Novais, Jake Fay, and Shaun Dougherty. Thank you to Erica Rosenfeld Halverson, who in addition to being my best friend for over 30 years, has also become an academic mentor. To the FW ladies, thank you for unconditional love and support. And to my girls, especially Sadie, Jackie, Emilie, and Sabina, thank you for being by my side always. Thank you K. Even in your absence, you encouraged me to push on and to take care of myself. I know you would have been proud.

Completing a dissertation while raising children was at times a complicated challenge. I am so thankful for my village, especially Colleen, Meaghan, Janna, Julie, Jed, Helen, Faye, and Debbie, all of whom at one point or another made sure my children were safe, happy, and occupied so that I could finish.

Finally, this dissertation would not have been possible without the support of my incredible family. My parents, Barbara and Stephen Litke, never stopped believing in me and never once questioned my ability to complete this program, even when I wasn't so sure. My in-laws Chuck and Estelle Bennett, never hesitated to help out and are the reason my family did not go hungry during these last few months. My boys, Oscar and Elliot, continue to amaze me daily and never stop reminding me about what matters in life. You inspire me and I love you both beyond words. My husband Nathan has been an incredible partner, parent, and advocate throughout this process. He weathered the ups and downs of this journey with me, patiently asking daily, "are you working tonight?" Here's hoping the answer to that question becomes "no" a lot more frequently.

Table of Contents

| | |
|---|-----|
| Abstract | vi |
| Chapter 1. Introduction | 1 |
| Chapter 2. The Format and Features of Contemporary Algebra Instruction: An Exploratory Study | 9 |
| Chapter 3. Measuring Instructional Practice in Algebra: The Development of An Observational Instrument | 70 |
| Chapter 4. The Nature and Quality of Algebra Instruction in Five Urban Districts | 107 |
| Appendices | 159 |
| References | 172 |
| Tables and Figures..... | 184 |
| Vita | 205 |

Abstract

Algebra is considered a linchpin for success in secondary mathematics, serving as a gatekeeper to higher-level courses. Access to algebra is also considered an important lever for educational equity. Yet despite its prominence, large-scale examinations of algebra instruction are rare. In my dissertation, I endeavor to better understand what contemporary algebra instruction looks like. I explore instructional practices across a large sample of video recorded algebra lessons from 5 urban districts. To do this, I draw on video and other data from the Measures of Effective Teaching (MET) Project. In the first study, I utilize grounded analysis to describe the format and features of instruction in lessons in the sample. I find that most lessons are teacher-centered with some opportunity for student engagement in mathematical thinking; however, very few lessons provide significant opportunities for student exploration or discovery of mathematical concepts. Looking beneath the surface, I find specific instructional practices teachers employ in algebra lessons and argue that improving these practices may be a promising lever for instructional improvement. Next, I describe the development and validation of an observational instrument oriented toward algebra and designed to measure the nature and quality of these practices. Finally, in the third paper, I use the observational instrument to describe the frequency and quality of these practices in algebra lessons in the sample. I present both descriptive results and qualitative cases of algebra lessons to illustrate these instructional features.

Chapter 1.

Introduction

We treat schools as doughnuts. We are very good at explaining the periphery... but we do not understand the hole in the center (what makes children learn). The center is the essence of schooling, where the fundamental but frequently unstated, priorities of education are generated, and our explicit understanding of that is a void (Graham, 2005, p. 198).

I was a teacher for six years in a New York City public school. I taught mathematics at a small, progressive 6th–12th grade school that was focused on inquiry-oriented instruction, employed project-based learning, and relied on portfolio assessment. I joined the staff as a new teacher and it was in this environment that I learned how to teach. I developed my teaching practice in a school culture that believed that all students could be successful in higher-level mathematics and that it was our job as teachers to ensure their success. Our students frequently came to us as middle schoolers performing below grade level and, by the time they entered my classroom as 12th graders, were taking Pre-Calculus.

My last year at the school, I taught an AP Calculus class to a small group of students who had accelerated through the high school mathematics sequence. These students were extraordinary in their ability to deeply understand the concepts of calculus, yet when they attempted to solve open response AP test questions, they faltered. While they could proficiently communicate the meaning of derivatives and anti-derivatives, they struggled to fluently execute the mathematical procedures necessary to enact their understanding and to flexibly call upon the appropriate strategy or knowledge in the appropriate context. This was an ever-present tension in my classroom that year and led me to wonder about how to bridge this gap. I began to ask questions: *What does it mean*

to understand complex mathematical concepts? How can teachers work with students to balance both understanding and procedural fluency? What does it mean to teach mathematics well? How might I improve as a teacher? This last question was particularly acute. I searched for advice, techniques, and even research oriented at instructional improvement in secondary mathematics and found little guidance. I encountered interesting tasks, general teaching strategies, and real world applications to be explored, but I was never able to find answers to the question of how I might improve the teaching of *this particular* content to *my particular* students.

Questions such as these have motivated my research. First, I am interested broadly in what instruction looks like in U. S. mathematics classrooms. How far are current classroom practices from where we might ideally want them to be? There has been important work done in this area at the elementary and middle school levels (see for example Hiebert, et al., 2005 and Hill, Litke, Lynch, Pollard, & Gilbert, 2014). The results do not show the student-centered and understanding-oriented classrooms that I was used to as a teacher to be the norm. While little research has explored instruction in large samples at the high school level, there is reason to believe that these instructional practices might be less likely to exist in contemporary high school classrooms (Banilower, et al., 2013; Powell, Farrar, & Cohen, 1985).

I focus the inquiry in this dissertation on algebra instruction specifically. I do so for both academic and personal reasons. Both researchers and policy makers have paid close attention to algebra in recent years. Proficiency in algebra has been touted as a national imperative (National Mathematics Advisory Panel, 2008), debated as a necessary evil (Hacker, 2012), and promoted as a civil-rights battleground (Gutstein, 2006; Moses

& Cobb, 2001). Hyperbole aside, it is a gatekeeper course, with successful completion a prerequisite for later courses in secondary mathematics (Kaput, 1995). My own teaching also revealed the importance of quality algebra instruction, as algebraic concepts and fluency were so crucial to my students' success in Calculus.

While improving access to this key course is an important and admirable goal, what do we know of the instruction that occurs inside the classroom once the gates are opened? It is this question that guides the papers in this dissertation. The broad question I am aiming to answer throughout is: *What does contemporary algebra instruction look like?* While a complete answer to this question is complex and multi-layered (and certainly outside the scope of this dissertation), I begin to chip away at an answer by engaging in a series of discrete steps. This includes 1) first determining *what* is currently happening in algebra classrooms by conducting a grounded analysis of extant instructional practices in videos of algebra lessons; 2) creating and validating an observation instrument to categorize that instruction; and then 3) applying it at scale—using the instrument to score a large sample of video instruction. These scores, alongside qualitative analyses of the lessons themselves, begin to paint a picture of contemporary algebra classes.

The studies in this dissertation are all grounded in a close analysis of classroom practice in algebra, specifically by watching (and re-watching) videos of algebra lessons. There are many possible aspects of instruction that I could have chosen as a focus for this inquiry, but I take as a theoretical framework the idea that the mathematical work that occurs in classrooms is distinct from classroom climate, pedagogical style, or the deployment of generic instructional strategies (Hill & Grossman, 2013). There are

certainly other important features of instruction—for example, classroom climate, management, teacher questions—that have been well-studied in the process-product literature. But I am focusing here on *mathematical* features of instruction, and particularly those features that may be salient to learning and understanding algebra. Next, I take the stance that instruction is comprised of the interaction between teachers, students, and content and that these interactions take place in classroom environments (Cohen, Raudenbush, & Ball, 2003). As such, I focus the inquiry largely at the lesson-level (rather than at the teacher-level), aiming for a close and careful analysis of the instruction therein.

All of the studies in this dissertation utilize data from the Measures of Effective Teaching (MET) Project, a Bill and Melinda Gates Foundation-funded study that aimed to determine fair and reliable ways to measure teacher quality. I use this data for a different purpose, however. The project partnered with approximately 3,000 4th–9th grade teachers across six urban districts: Charlotte-Mecklenburg, NC, Dallas, TX, Denver, CO, Hillsborough County, FL, Memphis, TN, and New York City, NY. Teachers on the project contributed up to four video-recorded lessons per year, making this one of the largest samples of classroom instruction available to researchers. As the ninth grade mathematics lessons were all from algebra classrooms, analysis of this data is primed to begin to address the questions articulated above.

The sample is not representative of urban districts, schools, and classrooms and as such, I do not aim to generalize the findings beyond the sample. But the findings *do* shed light on practices that appear salient across classrooms in this sample and that resonate with many previous descriptions of mathematics instruction in urban classrooms. It is my

hope that it also goes beyond confirming what we already know and believe to be true.

By uncovering more fine-grained instructional practices, I hope to provide further insight into teaching practices in algebra that will be of use to both researchers engaged in work with more representative samples, and to practitioners engaged in the important work of instructional improvement.

In what follows, I investigate the nature, frequency, and quality of particular instructional practices. In Chapter 2, I present an exploratory study of algebra instruction in 75 lessons from 24 ninth grade teachers drawn from the MET Project data. I conducted a grounded analysis in order to identify instructional practices that were salient in the sample and that might then be used in the development of an observational instrument specifically oriented towards algebra teaching. In the paper, I describe the themes that emerged from this analysis and illustrate each with examples from classroom practice.

In Chapter 3, I present the Quality of Instructional Practices in Algebra (QIPA), an observational instrument that was developed from the exploratory analysis conducted in Chapter 2. The QIPA protocol was designed to systematically describe the nature and quality of specific instructional practices in video-recorded algebra lessons, articulating what these practices look like across levels of quality. In this chapter, I describe the development of the instrument, present the domains and dimensions of the QIPA, and discuss the scoring procedures associated with the protocol. I piloted the protocol on a sample of 75 lessons and present a discussion of the quality of information provided by the instrument, focusing on issues of validity and reliability.

Finally, in Chapter 4, I present a descriptive study of algebra instruction in 108 lessons from a randomly selected sub-sample of 30 teachers from the larger MET sample.

I scored all lessons using the QIPA in order to further develop the picture of what contemporary algebra instruction looks like. I present results showing the frequency and quality of the instructional features captured by the observation instrument. To give a more nuanced illustration of instructional practice in algebra, I complement these results with case studies of two lessons: one that depicts “typical” practice in the sample and the other—a contrasting case—that illustrates high quality instruction on a number of practices in the observational protocol.

So what *does* algebra instruction look like? The findings of the analysis presented in this dissertation paint a picture of both stasis *and* change, but also of the need to focus on content- and discipline-specific practices. In some respects, the findings indicate that algebra classrooms in this sample in 2010 bear a striking resemblance to algebra classrooms in 1990 (and to the algebra classrooms of 1970). The format of instruction is largely teacher-led and, although there are opportunities for students to engage in the mathematics through solving problems and interacting with the teacher and one another, there is little inquiry or exploration into mathematical concepts. Despite decades of reform efforts by the mathematics education community, there is little engagement in highly cognitively demanding tasks, (productive) mathematical struggle, or mathematical discourse.

There is a long tradition of reforms failing to make inroads into classroom practice. Powell, Farrar and Cohen (1985) argue that, among other reasons, reformers’ directives to reform instruction frequently do not include the assistance to teachers necessary to actually change classroom practice. Importantly, it is also likely that teachers themselves did not experience reform-oriented instruction as students. Absent this

understanding, teachers continue to engage in familiar and comfortable practices. Cohen (2011) points to the ways in which the structures and organization of schooling in the U.S. work against instructional improvement. The structure and workload of high schools in particular may also work against reforming practice (Powell, et al., 1985). High school teachers frequently have large numbers of students across multiple class sections and might teach multiple courses (and grade-levels) in a given semester or year.

Yet instruction *has* changed in some ways and the results of this study do not simply confirm what others have found regarding the format of secondary instruction. Indeed, looking at a smaller grain size, I find that teachers *are* engaging in practices that reformers hope to see, at least to a modest degree. These encouraging features in teachers' extant instruction—what I call “glimmers” of promising practices—can be used as a foundation upon which to build and strengthen instruction. For example, algebra lessons frequently included multiple representations of similar algebraic ideas and problems (e.g. graphs, tables, equations, and contexts), although less frequently did they include explicit connections across representations at high levels and rarely did they include a discussion of how different representations reveal different information about the algebra concept under study. Better understanding these nuances and having language to ground discussions of deepening practices such as this presents an opportunity for those engaged in instructional improvement efforts.

The studies in this dissertation also broaden the discussion of what instructional features might matter in algebra, particularly in the context of teaching procedures. In the push for conceptual understanding, the teaching of procedures has been frequently disparaged, but procedures are ubiquitous in algebra and procedural knowledge is an

important element of algebraic understanding (Kieran, 2013; Star et al., 2015). One of the findings of this dissertation is that teachers engage in multiple instructional features *when* teaching procedures. For example, teachers engage in supporting students' flexibility in selecting between procedures, comparing and contrasting different procedures for efficiency, and highlighting for students key decision points within procedures. These practices have been shown in some cases to build students' procedural fluency (Star, et al., 2015). While these features did not appear at high levels of quality in most lessons, they provide an additional possible insight for instructional improvement efforts and a fertile ground for future research.

The problem of instructional improvement has long been a puzzle for scholars, policy makers, administrators, and teachers. Graham (2005) argues that policy interventions frequently treat schools as “doughnuts,” focusing on factors that influence the periphery (structure and organization, standards, curriculum, etc.) while leaving unexamined the “center” of the doughnut—the instruction that teachers engage in and that students experience in the classroom. This dissertation peers into the center of the doughnut, adding to the body of descriptive work around instructional practice in mathematics in the historical moment at which the ambitious Common Core State Standards are beginning to be adopted by schools nationwide. As schools and teachers wrestle with Common Core implementation, it is important to understand the degree to which instruction is aligned with expectation. The results of this study illuminate potentially important gaps between what the standards expect of teachers (and students) and the instruction that occurs in classrooms. But it also provides insight into particular instructional features that may be leverage points for instructional improvement.

Chapter 2.

The Format and Features of Contemporary Algebra Instruction: An Exploratory Study

There are few high school courses more present in the public discourse than algebra. From a policy perspective, success in algebra is thought to serve as a gatekeeper to higher-level mathematics (Stein, Kaufman, Sherman, & Hillen, 2011) and a predictor of later academic success (Adelman, 2006), resulting in an emphasis on access to this key course and the concepts it encompasses. Concurrently, policy makers and professional organizations have called for substantial shifts both in how teachers conceptualize algebra and in what they emphasize in the classroom (National Council of Teachers of Mathematics [NCTM] 1989, 2000). The resounding message over the past few decades has been that algebra instruction should work to jointly foster conceptual understanding, procedural fluency, and problem solving (Hiebert, 2003; NCTM, 1989, 2000), a stance echoed, if not amplified, by the Common Core State Standards for Mathematics and Standards for Mathematical Practices (National Governors Association Center for Best Practices [NGA], Council of Chief State School Officers [CCSSO], 2010). In its position statement in support of the Common Core, the National Council of Teachers of Mathematics argues that these new standards promote “the development of more rigorous, focused, and coherent mathematics curricula, instruction, and assessments that promote conceptual understanding and reasoning as well as skill fluency” (NCTM, 2013, par. 1). Setting the instructional bar as one that requires deep mathematical engagement from students has implications for both what is taught in algebra classes and how teachers interact with students around the content.

Despite the rhetoric around hoped-for instructional practices, little research has examined the nature of instruction in algebra classrooms. While mathematics education researchers argue for teaching practices that promote mathematical meaning-making and reasoning, we know comparatively little empirically about the specific features of instruction prevalent in contemporary algebra classrooms. Studies of classroom practice in algebra are usually small-scale and are frequently self-reflective—focusing on one or a small number of teachers (e.g. Chazan, 2000; Raymond & Leinenbach, 2000) in hopes of understanding specific practices associated with reform efforts (e.g. Spillane & Zeuli, 1999). As a consequence, discussion of the broader *qualities* of instruction present across classrooms is largely absent. For example, we know little about whether the recent waves of reform in mathematics appear in algebra classrooms or whether traditional modes of instruction persist. As schools and districts transition to the ambitious Common Core State Standards and attend to its instructional implications, it is fruitful to closely examine contemporary instruction. If we expect teachers to improve their instruction along these lines, we first need a better understanding of the nature of current practices. This understanding gives important perspective on the distance between where we are and where we hope to be, and what particular existing practices might be leveraged for improving student understanding and achievement.

In this paper, I capitalize on the rich video data available from the Measures of Effective Teaching (MET) Project, a three-year large-scale study of 3,000 teachers from six urban districts. I conducted a grounded analysis of video from a sample of ninth grade classrooms in order to identify themes and practices common to algebra teaching. While not a nationally representative sample, the MET project provides one of the largest

collections of recorded classroom instruction that exists, and as such provides an extraordinary opportunity to unearth themes across a large sample of lessons.

In what follows, I first frame the current study around issues in contemporary algebra and what is already known about the nature of algebra instruction. I next explain the data and methods employed in this exploratory study. I then identify the themes that emerged from the analysis, focusing on the content that is taught, the overall format of instruction, and the finer-grained instructional features that emerged in this sample. Finally, I discuss the implications of this picture for the reform goals articulated in the policy and research communities, and address the potential implications of these findings for future thinking around instructional improvement, particularly in the context of the preparation and development of teachers.

Background: Teaching and Learning in Algebra

Shifts in Contemporary School Algebra: Goals, Emphasis, and Expectations

In recent years, researchers and policy makers have focused their attention on school algebra, particularly in response to the widely accepted belief that algebra serves as a gateway to higher-level mathematics (National Mathematics Advisory Panel, 2008; Rakes, Valentine, McGatha, & Ronau, 2010; Stein, et al., 2011). Algebra is the bottom of the “layer cake” of high school mathematics (Kaput, 1995), with success in the course predicting students’ opportunities to pursue and succeed at higher-level mathematics (Schoenfeld, 1995). Furthermore, there is evidence that success in high school mathematics, particularly advanced courses, is correlated with positive outcomes such as college completion and persistence (Adelman, 2006). For these reasons, access to algebra has been trumpeted as an equity issue (Guiton & Oakes, 1995; Moses & Cobb, 2001;

Oakes, 1990), with many calling for opening the gates of algebra to more students, particularly those who have historically been underrepresented in both Algebra I classes and higher levels of mathematics.

These factors together have resulted in large increases in enrollment in school algebra across demographic groups (Chazan, 2008) and corresponding variability in the level of preparation students bring with them to algebra classes. These changes have been accompanied by the standards movement, which sought to shift the focus of both what is taught in algebra classes and how that instruction is delivered (Kilpatrick & Izak, 2008; NCTM, 1989, 2000, 2014; NGA, 2010). As a result, there have been changes in both what is emphasized in algebra classrooms, as well as in what students are expected to know and be able to do algebraically. These shifts further imply necessary changes in teachers' instructional practice.

While there is some consensus on the mathematical topics introduced and reviewed in an algebra courses (e.g. linear equations or solving single-variable equations), the curricular emphasis has shifted over time. The understanding of what algebra should entail has expanded beyond simply “generalized arithmetic”—an extrapolation of rules and computation with numbers—to include an emphasis on looking at patterns and structure, an emphasis on functions and relationships, and situating algebraic concepts in mathematical modeling of real-life contexts (Kieran, 2007). Whereas school algebra was once considered a method for solving particular types of problems, it is now also considered a means of describing and analyzing relationships (Usiskin, 1988). Indeed, the National Council of Teachers of Mathematics recently called algebra “a way of thinking

and a set of concepts that enable students to generalize, model, and analyze mathematical situations” (NCTM, 2008, par. 1).

These philosophical shifts have occurred alongside more concrete changes to the standards, and curriculum, and to the intended focus of classroom activities. Recent standards have emphasized classroom practices focused on meaning-making and conceptual understanding of mathematics (e.g. NCTM, 2000, 2013; NGA, 2010). In addition, the new Common Core State Standards for Mathematical Practice require that students reason, explain, construct mathematical arguments, model, and discern patterns (NGA, 2010), among other practices. The Common Core content standards for high school algebra also frequently require that students not only master particular skills and knowledge, but also that they demonstrate this mastery by explaining, interpreting, and making connections. In this environment, the typical problems in an algebra class have changed as well. Whereas many curricular tasks at one point asked students to “solve,” “expand,” and “simplify”—both in the context of symbolic expressions and equations and word problems—modern algebra curricula additionally ask students to “explain,” “predict,” “sketch,” “investigate,” or “explore” (Star, Herbel-Eisenmann, & Smith, 2000). This is frequently done in the context of applied (or “real-world”) problems (Kieran, 2007). Reform-oriented curricula that reflect these shifts often include a de-emphasis on paper-and-pencil computations.

Concurrent with these changes has been a focus on students demonstrating a conceptual understanding of algebra. Researchers and policy-makers alike advocate for algebra classrooms that are oriented towards meaning-making and reasoning in which students are engaged in productive mathematical struggle and high cognitive demand

tasks (e.g. Chazan, 2008; Hiebert, 2003; Hiebert & Grouws, 2007; Stacey & Chick, 2004), while simultaneously, developing procedural fluency (Kieran, 2007; Smith, 2014). The degree to which mathematics instruction, and algebra in particular, should focus on concepts or procedures (or both) is a disagreement that has raged for some time (for a discussion of this debate, see Hiebert, 2003; Kieran, 2013; Rittle-Johnson, Siegler, & Alibali, 2001). In fact, research shows that conceptually-oriented instruction had a greater impact on student achievement than instruction that focused solely on procedures (Rakes, et al., 2010). Yet most people in the mathematics education community acknowledge that the two types of knowledge are intertwined and likely develop in tandem (Kieran, 2007; Rittle-Johnson, et al., 2001).

This orientation towards developing conceptual understanding in addition to procedural fluency has resulted in changing demands around the structure of instruction in algebra classrooms. Some in the mathematics education community have advocated for instruction in which the teacher plays the role of facilitator and in which students engage in mathematical exploration in highly cognitively demanding activities (e.g. Stein & Lane, 1996). Others argue that students must be offered significant opportunity to engage in mathematical discussion and explanation (e.g. Chazan, 2000). Finally, some insist on classroom instruction rooted in using algebraic concepts to model real-world phenomena (Smith, 2014). This is a large departure from the traditional instructional format in which teachers deliver demonstrations of procedures and algorithms and then provide time for students to engage in independent practice (Hiebert, 2013). While there is evidence that these reform-oriented visions of instruction exist in individual schools and algebra classrooms (e.g. Boaler, 2002; Chazan, 2000), it is less clear—despite the

large shifts in curricular emphasis and standards outlined above—that large-scale changes in instructional format have occurred.

The Contemporary Algebra Classroom: Format and Features of Instruction

Empirical investigations into mathematics classrooms over the past 25 years imply that the algebra classroom of today may in fact bear a striking similarity to the mathematics classrooms from two generations prior, despite the shifts in emphasis, standards, and curriculum outlined above (Hiebert, 2013). There is some evidence that notwithstanding the evolution in goals and orientation, the *format of instruction* in many classrooms remains relatively unchanged. By format, I am referring to the overall lesson structures and characteristics such as the degree to which the instruction is teacher-led, and the amount and nature of the mathematical work done by the students. Prior research indicates that mathematics classrooms frequently follow variants of the “acquisition-application” format (Hiebert, 2013, p. 17). In this format, the instruction is largely teacher-driven, where teachers generally review homework, present new content to the class, apply that content to examples, and then ask students to practice with similar problems independently in groups (Cuoco, Goldenberg, & Mark, 1996; Hiebert, 2013; Hiebert & Grouws, 2007; Star, et al., 2000). While some reform-oriented curricula aimed to replace this format with investigation and exploration (Star, et al., 2000), curriculum does not necessarily dictate instructional practice (Stein, et al., 2011) and large-scale investigations of US mathematics classrooms consistently find procedurally-oriented, fragmented, and cognitively unchallenging lessons that follow this same general format (Hiebert, 2013). For example, the TIMSS study, a large-scale comparative study of eighth grade mathematics classrooms, found that in the U.S., instruction was highly procedural

with little opportunity for student investigation of mathematical concepts (Hiebert, et al., 2005; Jacobs, et al., 2006; Stigler & Hiebert, 1999). In addition, instruction was marked by an absence of student mathematical reasoning. In a more contemporary study, using similar data from the Measures of Effective Teaching project as the current analysis, Litke (2014) found that in eighth and ninth grade algebra lessons, there was little opportunity for students to participate and engage with mathematical content in cognitively activating ways—e.g asking mathematically motivated questions or engaging in high cognitive demand tasks.

These results might be interpreted as a failure at reforming instructional practice. However, it is possible that there is more progress than these results may indicate and a focus on the format of instruction may not tell the full story. A more fine-grained look at the *features of instruction*—the particular moment-to-moment instructional strategies teachers engage in when working with students around mathematical content regardless of instructional format—is necessary and may reveal a more optimistic picture. For example, one such feature prominent in the mathematics education literature is the use of and connections between multiple representations of algebraic ideas (Knuth, 2000). During instruction, teachers may utilize an equation, a graph, and a table to represent a linear function and engage in explicit connections between and among these representations. Another such feature is making sense of an algebraic procedure by giving mathematical meaning to its steps (Kieran, 2013). These features occur in smaller instructional moments and can occur in a variety of instructional formats.

In elementary mathematics, some research finds that despite little change in classroom format, there have been some incremental changes in instructional features

(Hill, et al., 2014). For instance, a recent study of 4th and 5th grade mathematics lessons from five urban districts found evidence of specific practices aimed at meaning-making, though these were in the context of largely teacher-centered instruction (Hill, et al., 2014). Teachers employed such practices as explanations for why procedures work, connections between mathematical representations and ideas, and comparisons of multiple solution methods. However, these features were most often instantiated only briefly or without much depth. Yet other contemporary studies have found few such features in U.S. classrooms. For instance, the Measures of Effective Teaching project, from which the sample for this study is drawn, found little evidence of mathematical sense-making or teachers using student ideas in elementary mathematics lessons (Kane & Staiger, 2012).

It is not clear whether these modest instances of positive instructional features seen at the elementary level occur in secondary classrooms. The TIMSS study, conducted in eighth grade classrooms, found that in addition to instruction being highly focused on teaching procedures, teachers seldom engaged in instruction that made meaning of procedures, seldom made connections between mathematical concepts, and rarely engaged in making connections between multiple representations (Hiebert, et al., 2005; Jacobs, et al., 2006; Stigler & Hiebert, 1999). A more contemporary analysis of 571 eighth and ninth grade algebra lessons from the MET project using project-generated scores on observational instruments found little evidence of features contributing to the depth of the mathematics offered to students such as mathematical explanations or connections made across mathematical representations (Litke, 2014).

Considering both format and features together, there is reason to believe that established instructional practices in high school may be even more difficult to disrupt

than in the early grades and the incremental changes seen at the elementary level may not have permeated high school classrooms. In analyzing eighth and ninth grade algebra lessons in the MET data, Litke (2014) finds that where reform-oriented practices exist, they are more likely to be in eighth grade lessons than in ninth grade lessons. This suggests that high school algebra instruction may be more traditional in nature than even eighth grade algebra instruction. One reason that reform practices may be slower to emerge in high school classrooms is what Grossman and Stodolsky (1995) call a shared subject-specific subculture at the high school level. They argue that high school teachers' perceptions of reform are shaped by their subject-specific beliefs. For mathematics teachers, concerns around coverage and pacing, students' mathematical preparedness, as well as beliefs regarding the sequential nature of mathematics may influence decisions around the adoption of hoped-for reforms. In addition, the structure of high schools may inhibit the adoption of more ambitious practice (Powell, et al., 1985). Many high school teachers teach multiple sections, amounting to over 100 students daily. In addition, high school teachers frequently have multiple "preparations," in which they teach multiple distinct courses in a given year (e.g. one section of algebra, one section of geometry, etc.). This organization of the work of teaching high school may contribute to the difficulty in incorporating reform-oriented practices.

Another view is that there may be something particular to algebra that distinguishes it from elementary mathematics, making it more challenging for students and thus more challenging to teach. For example, algebra represents a significant shift in the structure and content of school mathematics (Kieran, 1992) toward the abstract and symbolic (Star & Rittle-Johnson, 2009). Algebra curricula rely heavily on the use,

manipulation, and understanding of symbols, equations and expressions. While elementary mathematics focuses largely on finding numerical solutions to problems, in algebra, the attention shifts to understanding relationships.

Regardless of the reason, existing research indicates that instructional change in middle and secondary algebra classrooms may be slow to take root. But this view of instruction may not tell the whole story. First, existing observation instruments may be well-primed to capture instructional formats, but may miss more nuanced instructional features. Any observational instrument must take a particular theoretical and practical lens with which to view instruction. Doing so necessarily prioritizes some aspects of instruction over others. For example, the version of the Mathematical Quality of Instruction instrument used in the MET Project studies focused on larger-grained categories (e.g. the depth of the mathematics offered to students and the nature of the student engagement with the mathematics), but was less keyed to specific features of instruction (Kane & Staiger, 2012). Other protocols, such as those used in the TIMSS study, focused largely on the format of instruction—prioritizing the structure of lessons, the content of the mathematics offered to students and the nature of that content, how students worked on mathematics during the course of the lesson, and the kinds of mathematical reasoning engaged in by students (Hiebert, et al., 2005). This was a useful lens for the purposes of a comparative study, but did not focus as closely on instructional features.

In addition, to the degree that the observation instruments used in the studies outlined above do focus on instructional features, they are keyed to general strategies for all mathematics classrooms. As they are not specific to algebra teaching, they may miss

other key features present in secondary mathematics classrooms, and specifically in algebra. Looking closely at algebra instruction without a pre-determined observational rubric may allow for a better understanding of some of these practices. In other words, are there perhaps promising practices in algebra classrooms that extant instruments are not primed to capture? Looking at the mathematics education literature indicates that there are particular instructional practices that, while not unique to algebra, may be particularly salient or important in algebra instruction.

Potential Instructional Features Salient to Algebra Classrooms

The research literature in mathematics education gives some indication of potentially promising instructional features in algebra specifically (and in mathematics more generally) that we might hope to see in a close investigation of contemporary algebra lessons. Given the prevalence of teaching procedures in algebra, as well as the emphasis on “teaching for understanding” (Kilpatrick & Izak, 2008), researchers have begun to theorize around the nature of instruction on procedures. Theoretically, if students understand “how and why” a procedure works, they will be better able to adapt procedures flexibly to new situations (Hiebert, 2003, p. 17). Teaching focused on procedural skills at low levels of cognitive demand (such as recall or rote practice of well-established procedures) may ultimately hinder the development of conceptual understanding (Rakes, et al., 2010), but instruction that integrates concepts into the teaching of procedures may have promising effects. Indeed, the teaching of procedures can be either deep or superficial (Star, 2005) and it may be that high quality instruction on procedures with an attention to the concepts underpinning them improves conceptual understanding. Hiebert and Grouws (2007) suggest that teachers attend explicitly to

concepts in their instruction, a practice that could be instantiated regardless of instructional format. Concepts are defined here as “connections among facts, procedures, and ideas” (p. 383). Instructional features that explicitly and publically attend to these connections thus have the potential to promote conceptual understanding.

Another instructional feature advocated for by the mathematics education community includes capitalizing on connections among and between mathematical topics and representations. For example, Hiebert and Grouws (2007) advocate coherence and connections between mathematical topics. Another promising strategy is linking symbolic representations and abstract algebraic ideas to their numerical analogs (Kieran, 2007). In making explicit connections between abstract ideas and their numerical underpinnings, teachers theoretically develop students’ ability to develop linkages to mathematical properties they already know and understand.

Others suggest allowing students to struggle with important mathematics (Hiebert & Grouws, 2007). Struggle affords students the opportunity to wrestle with mathematical ideas. Productive mathematical struggle can occur when students are presented with tasks that are of high cognitive demand and are allowed to engage in discourse and analysis around the mathematics. It is important to note that, like attending to mathematical concepts, productive mathematical struggle can occur in multiple formats of instruction, from direct instruction to more student-centered discovery.

The features described above are generally under-studied in secondary classroom settings and it is unclear whether and to what degree they have taken root. Yet empirical research has shown some of these strategies to be promising levers for improving students’ skill and understanding. For example, research on comparing and contrasting

multiple solution methods to a given problem for similarities, differences, and efficiency can promote both procedural fluency and conceptual understanding (Lynch & Star, 2014b; Rittle-Johnson & Star, 2009). Instruction that focuses on form and structure in algebraic objects and relationships can also help to unearth conceptual underpinnings (Kieran; 2013). Others have shown the importance of students providing explanations, justifying their reasoning, and engaging in mathematical discussions (O'Connor, 1998; Star, et al., 2000). Finally there has been promising evidence for making connections across mathematical representations (e.g. graph to table to equation), as in order to move fluently between representations, students must understand the connection between them (Knuth, 2000). Linking multiple representations of algebraic ideas and procedures can also help students develop deeper understanding of equations and expressions (Chazan & Yerulshlamy, 2003). Research shows that students may need explicit assistance linking across representations (Chazan & Yerulshlamy, 2003), but little time in classrooms is spent on developing these connections (Knuth, 2000).

Responding to these concerns, many in the mathematics education community have argued for the need to emphasize “teaching for understanding” (Kilpatrick & Izak, 2008), advocating methods such as cooperative group learning, problem-based learning, and student inquiry (Rakes, et al., 2010), as well as practices such as using multiple solution methods and strategies, making connections between multiple representations (e.g. tables, graphs, and equations), and the increased use of real-world, contextualized problems (Kieran, 2007; NCTM, 1989, 2000; NGA, 2010). Yet it is unclear whether and to what degree the teaching of algebra has evolved along these lines.

A Fine-Grained Look at Contemporary Algebra Instruction

Despite the theoretical promise of the instructional features outlined above to promote student understanding in algebra, research that articulates the nature of instruction in algebra classrooms and investigates the practices therein is scarce. More common is research on teacher quality in mathematics that focuses on the relationship between teacher characteristics and student outcomes (Wayne & Youngs, 2003) or evaluates the degree to which teachers contribute to student achievement through value-added analyses (e.g. Kane & Staiger, 2012). We know too that many of these instructional features have permeated curriculum and have been a focus of professional development efforts in recent decades (BaniLower, et al., 2013), but it is as yet unknown whether and to what degree they have permeated classrooms. As researchers and policy makers remain concerned about gaps in mathematics achievement among students, particularly in urban schools (Lankford, Loeb, & Wyckoff, 2002), investigating the nature of instruction in contemporary algebra classrooms in urban districts may be particularly important. Looking closely at the nature of algebra instruction in such classrooms will help inform research and efforts at instructional improvement in these contexts.

Research Questions

It is clear that more research is necessary to understand the nature of instruction in algebra classrooms and to identify algebra-specific features of instruction. If we view mathematics reform as moving from teacher-centered practices to inquiry-oriented ones, we may find that little has changed in many secondary algebra classrooms. However, if we deepen the analysis to uncover instructional features prevalent in contemporary

algebra classrooms, we may see a more optimistic picture. To address this, I conducted a grounded analysis of a large sample of ninth grade algebra lessons, asking the following:

- *What is the format of algebra instruction in a sample of ninth grade classrooms from five urban districts?*
- *What are key instructional features in this sample of ninth grade algebra lessons?*

Research Design

Data

Data for this study comes from the Measures of Effective Teaching (MET) Project. The project partnered with approximately 3,000 teachers across six urban districts: Charlotte-Mecklenburg, NC, Dallas, TX, Denver, CO, Hillsborough County, FL, Memphis, TN, and New York City, NY. Teachers on the project contributed up to four video recorded lessons per year over the 2009–2011 schools years. MET Project raters scored these videos on a number of classroom observation instruments, both general and subject-specific. The project also collected teacher demographic information, administered assessments of teachers' content knowledge for teaching and student's mathematical knowledge (e.g. SAT-9), and retrieved student achievement and demographic data from the partner districts. In addition, the project administered student, teacher, and administrator surveys. The overall goal of the MET project was to determine fair and reliable methods for measuring effective teaching (Kane & Staiger, 2012).¹

Data Sources

Classroom video. This study relies heavily on the video-recorded classroom lessons from ninth grade teachers. Video-recorded classroom lessons allow for a deep

¹ For more information on the design and components of the MET Project, see www.metproject.org.

analysis of the features of classroom instruction. Video also permits researchers to slow down and re-watch classroom moments that may be missed in live observation (Hamre, Pianta, Mashburn, & Downer, 2007; Learning Mathematics for Teaching Project [LMTP], 2011). Video-recorded lessons allow for iterative analytic processes in which themes that emerge from the video are refined and applied to new lessons (Jacobs, Kawanaka, & Stigler, 1999). Lessons were video-recorded using an un-manned panoramic digital video camera operated remotely by teachers or school personnel (Bill & Melinda Gates Foundation, 2010a). While teachers selected the lessons to be recorded, the project requested that at least two of the lessons be focused on core topics in the teacher's subject area. Teachers uploaded their videos to a secure website, where they identified the topic of the lesson and had the option to upload supporting materials.

Content Knowledge for Teaching: Algebra assessment. The MET Project administered an assessment of pedagogical content knowledge to all teachers in the study. Those who taught ninth grade algebra took a specially-designed Content Knowledge for Teaching Algebra assessment (CKT) developed by researchers at the Educational Testing Service (ETS) and the University of Michigan (Bill & Melinda Gates Foundation, 2010b; Gitomer, Phelps, Wren, Howell, & Croft, 2014). The CKT assessment was administered in early 2011 and focused on teachers' pedagogical content knowledge—the specialized knowledge that enables teachers to effectively teach algebra content to students.² The Algebra CKT exam included 37 selected response items (two of which were excluded for poor performance) and had an overall reliability of 0.77 (Gitomer, Phelps, et al., 2014).³

² For more information on the CKT assessments developed for MET see Gitomer, Phelps, et al. (2014).

³ Although this reliability is lower than traditional tests of teacher knowledge, it is important to note that the MET CKT assessment was shorter in length, and researchers calculated that doubling the length of the test would yield reliabilities of between 0.82 and 0.91 (Gitomer, Phelps, et al., 2014).

Demographic data. The MET Project data includes demographic and background information at the teacher level. This information includes teachers' gender, race, years of experience, and years of experience in the district, as well as whether teachers have a master's degree.

Sample

For this study, I focused on the available ninth grade mathematics classroom videos from the first year of the MET study. In particular, I focused on the subsample of ninth grade algebra lessons from 233 teachers across five districts.⁴ I excluded teachers with no scores on the Content Knowledge for Teaching (CKT) Algebra assessment administered by MET (92 teachers) and teachers with no viewable videos (60 additional teachers), as both criteria were important to my sampling and analytic strategy (see below).⁵ My final analytic sample consisted of 81 ninth grade math teachers from 49 schools spread across five of the partner districts. Teachers have between one and five viewable videos each—with most teachers having four videos—yielding a total of 292 video-recorded algebra lessons.

The analytic sample is largely similar in demographic composition to the full ninth grade sample, with some minor differences. The analytic sample is comprised of more white teachers (65% compared to 53%), fewer Black teachers (22% compared to 28%), and fewer teachers who are not White, Black or Hispanic (1% as compared to 6%). Analytic sample teachers for whom there is information on experience and education have been teaching for slightly fewer years on average (6 years compared to 7.5 years)

⁴ While the larger MET study includes six partner districts, one district did not include any ninth grade math teachers as project participants therefore this study is restricted to five districts. See Appendix A for a breakdown of district representation in the full sample, the ninth grade sample, and the analytic sample.

⁵ While all teachers contributed video to the main MET study, approximately one third of teachers did not consent to make their videos available to researchers at the conclusion of the study.

but are more likely to have a Master's Degree or higher (32% compared to 28%).

Comparing the sample of ninth grade teachers with CKT Scores (n=141) to those with CKT scores and viewable video (n=81), there are no discernable differences in mean, median or range of these scores. For a more detailed comparison of the analytic sample to the full ninth grade sample and the ninth grade sample with CKT scores, see Table 1.

From the sample of 81 ninth grade algebra teachers, I selected an analytic subsample of teachers from among the top, middle, and bottom quintiles of the distribution of teachers' scores on the CKT assessment.⁶ Past research has found that instructional quality varies sharply according to teachers' mathematical knowledge (Charalambous & Hill, 2012; Hill, Ball, & Schilling, 2008), and by sampling in this manner I hoped to maximize the variability in algebra instruction included in the analysis.⁷ I used CKT scores from the MET-administered teacher assessment to rank the ninth grade algebra teachers in the analytic sample (n=81), binning teachers into quintiles. I then chose a random stratified sample of 24 teachers, eight teachers each from the high (top quintile), mid (third quintile) and low (bottom quintile) "quality" bins.⁸ This sampling yielded a total of 75 lessons with viewable video. The results of a study of instructional practice derived from this sample will provide insight into themes and instructional practices present across a range of lessons, however these themes are not intended to generalize to all ninth grade algebra classrooms. Rather, the results of this

⁶ Sampling procedures that use the MET's extant observational instrument scores to select teachers and lessons might result in a sample of instruction that privileges the factors examined by those protocols. Assuming teaching quality is normally distributed, taking a simple random sample from eligible teachers may have resulted in an oversample of typical lessons, thus underestimating variability among teachers and obscuring important instructional practices in the tails of the distribution. Prior work with the 5-point Mathematical Quality of Instruction instrument has shown observational scores to be normally distributed.

⁷ CKT scores in the sample are approximately normally distributed.

⁸ It is important to acknowledge that effective *teachers* and effective *teaching* are potentially distinct and that using teacher-level scores may mask variability in teaching practices. However, metrics such as this are commonly used as a proxy for teaching effectiveness.

exploratory analysis can be used as the foundation for further investigation into a more representative sample of classrooms.

Data Analysis

The goal of this exploratory analysis was to understand instructional formats and identify salient instructional features across a large sample of algebra lessons, developing descriptions of instructional practice specifically attuned to algebra. To do so, a research group of four experienced math educators engaged in an iterative process of watching video, discussing instructional features, returning to the literature, and re-watching video.⁹ We began by reviewing the existing instruments used in the MET study (e.g. MQI and CLASS) and the results of MET reports on instructional practice (e.g. Kane & Staiger, 2012). We also reviewed the literature on desired and effective practices in algebra to broadly understand what themes we might expect to see. Next, we began watching video from the subsample, blind to teachers' CKT scores. To do this, I first randomly selected a teacher from the subsample, then randomly selected a video from that teacher, and finally randomly assigned the video to a member of the research team. This process was repeated for each member of the research team.

Each researcher watched their assigned video in its entirety, following a protocol in which they recorded a brief lesson summary, including the topic of the lesson, a narrative description of the lesson, mathematical strengths and weaknesses of the lesson, and any salient instructional features. Each researcher then nominated key segments of

⁹ The research group was comprised of four former and current secondary mathematics teachers, two of whom were researchers in mathematics education, one of whom was a practicing algebra teacher and teacher-trainer, and one of whom was project staff on a research project on instructional quality in mathematics. Three of the members of the research group were trained raters on existing observational instruments (such as MQI and CLASS), one of whom had been involved in instrument development for the MET Project.

their assigned video that highlighted interesting practices, unique features of instruction, or (in later rounds) common features of instruction for the rating group to watch together and study in further detail.

In order to develop an understanding of the content of the mathematics and the instructional format among lessons in the sample, I relied on the researcher-generated lesson summaries. I first used these lesson summaries to analyze the algebra content covered in the lessons in this sample. Each lesson summary included the topic of the lesson as identified by the observer. I categorized the topics using standard Algebra I curricula and the Common Core State Standards. The lesson summary protocols also included a narrative description of each lesson. I coded these narrative descriptions for lesson format, using both established formats from the literature, and open coding to describe each lesson's format. I aggregated these codes to categorize the format of instruction, relying on researcher notes and, where necessary, research team discussion to clarify findings.

To determine instructional features, we relied on a thematic analysis using open coding via an iterative process (Corbin & Strauss, 2008). As stated above, researchers nominated key segments of each lesson they reviewed for the group to watch and discuss. At the onset, the identification of the instructional features was completely open—researchers were directed to choose whatever they found salient. As we progressed and focused our attention on particular themes, researchers nominated segments that reflected these practices across levels of quality in order to clarify the themes that emerged. The nominated segments formed the basis of the research team's meetings and discussions. For the nominated segments, each member of the research group first independently

watched the nominated segment, individually recording our observations on an observation protocol in which we described the segment topic, noted any salient instructional features, and commented on the particular features we noted. We did this for between two and four segments per week.

We compiled our observations, which were shared with the group, and we met weekly to discuss them, identifying and examining both convergent features and divergent ones. We then repeated this process with a new set of video segments selected from those nominated by the research group. I took notes during our meetings and wrote analytic memos after every meeting, summarizing the notes and synthesizing the research team's findings. After every four to five meetings, I coded our meeting notes for common themes and began to develop broader categories that captured these themes, writing memos articulating these themes and categories. We used these themes to inform future video segment selection. From this analysis, I generated a list of instructional features that I brought to the research team to discuss and refine. We next applied emerging instructional features to video, noting both positive and negative cases. We continued to watch video until saturation, where we were confident that no new themes emerged (Corbin & Strauss, 2008; Guest, Bunce, & Johnson, 2006). We identified saturation after watching 35 lessons, although we continued to watch video until we had seen 42 lessons in their entirety and discussed 38 segments.

To check the validity of the themes that emerged during group discussions, I also reviewed observer-written lesson summaries for all 42 lessons. I engaged in both a thematic and open coding process (Corbin & Strauss, 2008) in which I both applied codes developed from the literature and our research team to the data as well as developed

codes that emerged directly from the data itself. I then clustered these codes in thematic groups (Maxwell, 2013). I re-read the lesson summaries, coding them with the emergent themes and categories. Throughout, I took care to note both convergence and divergence with the results from the research group discussions. Through this process, I developed the formats and features I present below.

Findings

My analysis yielded results around both the format of instruction generally and specific instructional features that teachers employ. Below, I first ground the findings in a brief description of the nature of the mathematical content taught in lessons in this sample. Next, I turn to overall patterns across lessons in the format of instruction, noting the ways in which lessons conform to or diverge from the general trend. I then describe the particular instructional features that emerged from our analysis of the data, focusing on two broad categories: instruction on procedures and the leveraging of connections. For each category, I articulate the specific instructional features that arose in the analysis and illustrate these practices with examples from lessons in the data. Finally, I discuss other instructional practices that we might have expected to see given the literature but were not as salient in the data, noting examples of their frequency and quality.

Nature of the Algebra Content

As standards have shifted and more students have enrolled in algebra, some have raised concerns that the content of algebra classes has been watered down (Porter, Floden, & Fuhrman, 1998). To determine the degree to which lessons in this sample represented content that might be considered algebra, I recorded the main topic or objective of each lesson and aggregated these topics into broader categories. I compared the topics with the

Common Core State Standards, as well as standards from the states represented by the MET partner districts. Lessons in the sample covered a range of topics typically found in an eighth or ninth grade algebra class and reflected in the Common Core State Standards for Mathematics (NGA, 2010). For example, many lessons in the sample ($n=14$) focused on lines and linear relationships. Six lessons focused on single-variable linear equations through instruction on graphing lines or finding slope and intercepts. Another large cluster of lessons focused on solving systems of linear equations ($n=6$) or systems of linear inequalities ($n=4$). Lessons focused on symbolic manipulation also featured prominently ($n=9$), with five lessons featuring instruction on the rules governing exponentiation with variables with a common base, two lessons featuring the manipulation of radicals in equations, one lesson covering solving single-variable equations for an unknown, and one lesson focusing on multiplying binomials.

A smaller number of lessons ($n=6$) focused on functions and their properties: one lesson focused on the difference between functions and relations, two lessons looked at the properties of quadratic functions (e.g. finding the vertex and axis of symmetry), two lessons featured a focus on parabolic paths, and one lesson focused on the transformations of numerous parent graphs such as absolute value and cubic functions. Only a small number of lessons ($n=3$) were focused primarily on topics that would be considered part of a pre-algebra class (e.g. one introductory lesson on graphing in the coordinate plane, one a lesson on percent change, and one lesson on measures of central tendency). Another small group of lessons ($n=3$) had a more geometry-oriented focus (e.g. a lesson on angles and angle relationships, a lesson on the Pythagorean Theorem, and a lesson on trigonometric ratios), while another group of lessons ($n=3$) focused on

arithmetic and geometric sequences, a topic that is not necessarily considered part of a traditional Algebra I course, but is often taught in this context. Finally, two lessons were characterized as “mixed review,” in which the teacher led the class through a series of topically disconnected problems. In contrast to these two lessons, the rest of the lessons in the sample featured a more coherent focus on a single topic or small group of related topics.

Format of Instruction

I relied on observer-written lesson summaries to identify and categorize the format of each lesson. My coding of lesson summaries revealed that the archetypal mathematics lesson format prevalent in the literature—the teacher delivers content and models how to solve problems, and then students practice what they have learned—was reflected in lessons in this sample. In the majority of lessons ($n=34$), instruction was largely teacher-led with occasional opportunities for independent or group practice or brief opportunities for students to answer open-ended questions. In these lessons, the students participated in the lesson by occasionally providing brief but meaningful mathematical contributions (such as asking questions about the mathematics or providing mathematical explanations) and/or were given short periods of time to work through mathematical concepts on their own. Yet even with this student participation, the instruction was largely teacher-driven.

While these lessons were not always structured in exactly the same way, they frequently followed a variant of the acquisition–application format prevalent in the literature (Hiebert, 2003) in which the teacher introduced a new concept or talked through the procedure for a particular problem or problem type, followed by worked example

problems that the teacher did for or with the class. Students were then given some limited/timed opportunity for individual or group practice. In some cases, students presented solutions for these examples, while in other cases the teacher reviewed the answers. For example, in a lesson on graphing linear equations that was typical of this format, the teacher asked students to work silently on a warm-up problem for about five minutes. After giving students the solution to the problem, she demonstrated the steps for graphing a line in slope-intercept form on a Smart Board presentation that she revealed step by step.¹⁰ She directed the students to take notes on the steps. While she did this, some students offered occasional substantive mathematical contributions such as defining terms and asking brief mathematically motivated questions. Next, the teacher modeled how to graph two lines, after which students worked in groups of three to complete several practice problems. These problems were fairly straightforward graphing tasks that looked similar in structure to the problems the teacher had modeled. Finally, selected students were asked to write their solutions on the white board for the class. Many lessons followed this format of short presentation, worked example (either by the teacher or worked together with students), student practice, and review of student solutions.

While the majority of lessons followed this format, an additional four lessons departed from the archetype in that they were even more extreme in the degree to which the teacher directed the instruction. These lessons featured instruction that was entirely teacher-led with little or no student practice and limited student participation. What little student involvement there was in these lessons consisted largely of one-word responses or brief engagement in low cognitive demand activities such as short computation

¹⁰ In order to preserve teacher anonymity and the confidentiality of teachers in the data, I refer to all teachers in this study as female regardless of their actual gender.

exercises. The content of these lessons was usually presented in tightly constructed static PowerPoint presentations or Smart Board Notebook presentations. Material was either entirely written out in advance and revealed little by little with no opportunity for student input or presented with some opportunities for students to “fill in blanks” with words or solutions. In these lessons, observers noted a good deal of student passivity, with students having little opportunity to solve mathematical problems without assistance or step-by-step instructions. Instead, student participation was limited to copying notes from the pre-prepared teacher presentation.

For example, in one such lesson, after giving students a short bellwork assignment of two problems involving exponents and two questions about identifying the base and exponent in an expression, the teacher delivered a thirty-minute lecture articulating the seven properties of exponents (e.g. when multiplying expressions with a common base, you add the exponents) that were sequentially revealed on a pre-prepared Smart Board presentation. She instructed the students to copy what was written on the Smart Board into their notes. During her lecture, the teacher gave one example problem for each property, which she solved and the students also copied into their notes. She ended the lesson with one example of simplifying a rational expression with exponents that necessitated using multiple properties. She then modeled how to solve the problem using the properties. There was virtually no observable student participation in the lesson beyond students copying the teacher’s notes into their notebooks. In fact, the only student contributions throughout the entire lesson were brief responses to bounded questions such as, “What’s the name of that property?” “What is two squared?” and “What is four minus two?”

Given the focus and nature of the mathematics reform efforts, I expected to see some evidence in the sample of classrooms formats oriented around student understanding and meaning-making, through engaging in cognitively activating tasks (Stein, Smith, Henningsen, & Silver, 2000), tasks and activities involving student exploration or inquiry, students taking an active role in their learning (Kaput, 1999), or through solving multi-layered, contextualized, real-world examples (Boaler, 2002). Yet in this sample we saw little evidence of this mode of instruction. In fact, only two lessons were coded as having a significant portion of the lesson focused on student exploration and inquiry. In one of these lessons, students briefly worked in small groups to graph a contextualized situation in which an object followed a parabolic path. After going over this problem, the students were instructed to work on a mathematical investigation of a rocket launch. In this task, students were given a situation in which a rocket travels for a period of time according to a path described by a given parabolic equation. The task required that students answer open-ended questions about the rocket's trajectory, timing, height, and acceleration and that students graph the parabolic path. The task asked students to make sense of both their graph and the equation modeling the situation with a number of open-ended questions. The teacher introduced the task with little scaffolding, saying:

On page 497, they give you a description of the situation and give you this as the equation. What you're going to do is answer questions about what this equation means and what the numbers and variables in the equation mean. And then you're going to graph it. I don't want you to necessarily draw a graph on paper, but I do want you to make a graph on the graphing calculator.... We're going to do the entire investigation today and we're going to help each other.

Students worked on the task in groups, interpreting the questions posed by the task and asking the teacher for guidance when they were unsure of what to do. The teacher circulated around the room and, when she interacted with a group seeking her help, asked open-ended questions that encouraged student thinking, such as, "What is [the question] asking you?" "Looking at the equation, what are the variables telling you?" or "How is gravity represented in this equation? How does it show in this equation that gravity is making the rocket come down?" She also posed questions designed to ask students to explain their reasoning and uncover misconceptions. This continued for about 30 minutes, at which point the teacher brought the class back together to discuss some of the answers to the investigation questions (e.g. Approximately how long does it take for the rocket to reach a particular height on the way up and the way down?) using multiple representations to do so (e.g. tracing on the graph, looking at the table, using the equation). In this lesson, students were actively engaged in doing mathematics for the vast majority of the class period.

In the other lesson, although the focus of the lesson as a whole was less on student exploration, students were given a relatively open-ended contextualized problem to solve and worked on it at length. In contrast to the first example, however, this lesson also included an extended teacher-led presentation. The teacher first directed an extensive review on the procedure for graphing a system of inequalities. The teacher then read aloud a complicated contextualized problem involving gardening with a number of different kinds of plants with particular constraints around water, fertilizer, and sunlight intended to represent a situation that could be modeled with a system of linear inequalities. Students were put into groups to answer the questions presented in the task.

The teacher charged specific students with different parts of the task and the directions were heavily scaffolded (e.g. Partner A must graph the situation). Students worked on the task for approximately 20 minutes, after which the teacher asked a group of students to graph the system on the Smart Board at the front of the class. The teacher ended the class by going over the answers to the questions in the task (“Which type of plant can be placed on the dotted line?” “Can the rose live on the point (6,6)?”), asking students to explain their answers.

The remaining lessons in the sample ($n=2$) were review lessons (presumably for a unit or standardized assessment) and were structured around solving a series of problems and reviewing their solutions. Both lessons (from different teachers) were structurally similar. In one, the teacher displayed a sequence of questions on disparate algebra topics on a SmartBoard screen (e.g. simplifying an expression, finding slope, completing a sequence). For each question, the students were given a short amount of time to choose from multiple choice options and “buzz” their answer in using a computer system that recorded and displayed the class’ responses. After each question, the teacher announced the percent of students who answered the question correctly. She then either went over the solution to the problem, reminded students of “tricks” they could use to get the right answer, or in some cases simply moved on to the next problem. This activity comprised the entire lesson. While students completed a number of problems in these lessons, there was little explanation or discussion of the mathematics involved.

Features of Instruction

The findings above about the format of instruction are aligned with previous studies in the literature (e.g. Hiebert, et al., 2005). Yet upon a finer-grained examination,

a more nuanced picture emerged of specific instructional features that teachers engage in when teaching algebra. Analysis of lesson summaries, identification of features from nominated segments, and research team discussions and refinement revealed that these features clustered into two broader categories—Instruction on Procedures and Leveraging Connections to Build Understanding—each with three themes. In what follows, I first discuss the broader categories and then describe each of the major themes that emerged, illustrating each theme with examples from classroom lessons.

Instruction on procedures. Analysis of the lesson summaries showed that most of the lessons in the sample included significant amounts of instruction on procedures. This may not be surprising, as procedures form a large part of the content of the algebra curriculum (Kieran, 2013). In this sample, instruction on procedures happened in 38 of the 42 lessons. These lessons featured instruction on how to accomplish a particular skill or algorithm or how to employ a formula (e.g. how to solve a system of equations using elimination, how to solve a system of inequalities, how to find the vertex of a parabola, how to apply the rules of exponents to simplify expressions). We categorized instruction on procedures as occurring when teachers engaged directly in the presentation of new rules, formulas or algorithms, as well as when they reviewed previously learned procedures, or described procedures used in the context of solving problems. While teaching procedures was ubiquitous, the proportion of a given lesson spent on teaching procedures varied.

Given the prevalence of instruction on procedures in the sample and in the domain of algebra, the research team analyzed lesson video and lesson summaries to further investigate particular instructional practices teachers employed while teaching

procedures, focusing our analysis on the finer-grained qualities of teachers' instruction on procedures. While procedures were sometimes presented exclusively as a series of steps—as recipes to be followed—observers noted that at other times, there was more depth and nuance to the presentation of the procedures. In particular, our analysis of video-recorded lessons and segments of lessons yielded three features common in the teaching of procedures salient in this sample: the ways in which teachers and students made sense of procedures, the ways in which teachers supported procedural flexibility, and the degree of detail and organization in the presentation of the procedures.

Making sense of procedures. Consistent with contemporary thinking on the interweaving of concepts with procedures in algebra (Kieran, 2013; Star, 2005), instruction on procedures in this sample featured instances in which teachers worked to make sense of the procedures being presented. One way in which teachers did this was by attending to the meaning of individual steps of a procedure. For example, in teaching students the procedure for solving systems of equations by substitution, a teacher made meaning of the first step in the procedure, explaining why it was possible to replace the y in one equation with the right-hand expression from the second equation, emphasizing the idea and property of equivalence. Teachers also made meaning of the solution that resulted from procedures. For example, after going through the procedure for solving a system of equations using elimination, another teacher made sense of the solution generated by the procedure by underscoring that the values students found for x and y represented the x - and y -coordinates of the point of intersection of graphs represented by the two linear equations.

Another way in which teachers made sense of procedures was to attend to the purpose or mathematical goal of a procedure. For example, in teaching the quadratic formula, one teacher reminded students that the point of the algebraic manipulation was to yield the roots of a parabola. Teachers also attended to the mathematical properties underlying a procedure. For example, when teaching students how to multiply two binomials, a teacher underscored that the multiplication used in the “FOIL method” (a mnemonic reminding students the order in which to multiply terms—First, Outside, Inside, Last) was really a variant of the distributive property. In another lesson in which a teacher explained the procedure for adding exponents when multiplying expressions with a common base, she related the formula to the idea of repeated addition.

The length and quality of sense-making around procedures varied across lessons. In some lessons, teachers attended only briefly to making sense of procedures. For example, in a lesson on linear inequalities, the teacher first reviewed the procedure for graphing a system of inequalities. She narrated the procedure for students in a step-by-step fashion with little attention to meaning (e.g. first graph it as if it was a line, then decide if you have a dashed or solid line based on the inequality sign, then shade up or down). At the end of this segment of instruction however, she asked students if they remembered the procedure for testing whether the point $(0, 0)$ is in the solution set to a particular linear inequality. A student responded that they should plug the point into the inequality for x and y . The teacher did so, ending with the inequality $0 > 1$. The teacher then asked, “So what does that signify? That $0 > 1$. What is the significance of that?” She went on to briefly discuss why this inequality indicates that $(0, 0)$ is not part of the solution set, making meaning of this particular step in the procedure. While the majority

of the presentation of the procedure was not attuned to meaning, there was brief attention to sense-making around the use of a test point.

In some lessons, however, making sense of procedures featured prominently. In a different lesson on linear inequalities, for example, another teacher introduced the procedure for graphing inequalities with the example $x + y \leq 2$. She had the students graph the line $x + y = 2$ and introduced the idea of the shaded region by asking students whether particular points satisfied the inequality. Students began to notice that points below and to the left of the graphed line satisfied the inequality and continued to suggest points in that area. A student asked if they show that there are many points that make the inequality true by shading them. The teacher responded:

We're talking about not a line, but a whole region—an area where the points $x + y$ are less than two. As a matter of fact, *all* of the points below this line and to the left. . . all of these points [draws multiple points in the region] satisfy the inequality $x + y \leq 2$. This point down here is $(0, -7)$. Zero plus negative seven is less than two. So when we graph an inequality, we end up with a whole region or area and what we do is shade that area. We're not just talking about points on a line. We're talking about a whole region, or area, where all the points fit the inequality. And we shade that whole area.

In this exchange, the teacher made sense of the shaded region as the solution to the inequality. Next, the teacher turned to the boundary line of the shaded region at $x + y = 2$. She asked students whether the points on the line $x + y = 2$ were part of solution set for the inequality $x + y \leq 2$, engaging the students in a discussion of why this was true. She explained to the students that the inequality $x + y \leq 2$ is inclusive of the line $x + y = 2$ because it contains the equality, plugging in the point $(1, 1)$ to make this clear to students. She used this idea to explain why the inequality is graphed with a solid (as opposed to dashed) line, giving meaning to this step in the procedure. Indeed, throughout the

segment of instruction on how to graph linear inequalities, the teacher attended to the meaning of individual steps of the procedure and the solution generated by the procedure. Much has been made of the importance of interweaving the teaching of procedures and the teaching of concepts (Kieran 2013), and encouragingly in this sample, we saw teachers attend to concepts in the context of teaching procedure. Doing affords students the opportunity to understand how and why procedures hold (Star, 2005).

Despite its promise, we also saw lessons in which teachers taught procedures without engaging in any sense-making. In these lessons, procedures were taught as recipes to be followed and divorced from their mathematical meaning. For example, in one such lesson, a teacher introduced the procedure for solving quadratic equations using the zero product property. She explained to students that in order to solve for x , they needed to first set the equation equal to zero, then factor the quadratic, and then set each factor equal to zero and solve. While she was correct in her description of the procedure, she did not make sense of the procedure as a whole, or give meaning to any of the individual steps (e.g. *why* it is possible to set each factor equal to zero). Students copied down a model example and then were expected to reproduce the procedure on a number of practice examples.

Supporting procedural flexibility. In addition to developing conceptual understanding of algebra through the learning of procedures, deep procedural knowledge also encompasses student flexibility in the use and application of procedures (Star, 2005). In this sample, observers found that teachers engaged in practices that afforded students the opportunity to begin to develop procedural flexibility—knowing which procedure to

apply in a given situation, knowing when to apply it, and attending to key decision points in a procedure.

Observers noted teachers engaging in this practice in multiple ways. In some lessons, teachers noted multiple pathways through a procedure. For example, in simplifying rational expressions with radicals, one teacher commented that students could either rationalize the denominator at the outset of the problem or simplify other aspects of the expression first, and then attend to the radicals. Teachers also attended to when a particular procedure might be applicable in a given situation. For example, in a lesson on graphing linear equations, a teacher highlighted how the form of the equation might lead to the choice of a particular method for graphing. If the equation was in standard form for example ($Ax + By = C$), then students might consider finding the x - and y - intercepts by plugging in zero for x and then solving for y and vice versa. Alternatively, they might instead choose to manipulate the equation to be in slope-intercept form. This attention to flexibility may afford students the opportunity to see and understand algebraic structure (Star & Rittle-Johnson, 2009).

Teachers also supported procedural flexibility by attending to key conditions for steps within a procedure, noting what must hold in order for a particular step to be executed or when a procedural decision must be made. For example, in reviewing the procedure for solving a system of equations using elimination, a teacher emphasized the role of the sign of the coefficients in the two equations. She explained that when a variable appears in both equations with the same coefficients but with opposite signs (e.g. $+3x$ and $-3x$), it is possible to add the equations together and eliminate the variable. However, when *both* the coefficients and their signs are the same (e.g. $+3x$ and $+3x$),

students would need to make a choice to either subtract the two equations in order to eliminate the variable or multiply one equation by negative one and then add the equations together.

Finally, observers noted teachers supporting flexibility by comparing multiple procedures for their affordances and/or limitations. For example, after teaching students the procedure for solving a system of equations by substitution, one teacher reminded students that they had recently solved similar systems by graphing, noting that while graphing might be easier or more efficient for some problems, in cases where the coordinates of the point of intersection are not integers, the algebraic method may be a better way to proceed.

In some lessons, this practice featured prominently. For example, in a lesson on solving systems of equations using elimination, a teacher presented an example that required students to multiply one equation by a constant in order to be able to eliminate one variable (see Figure 1). The teacher demonstrated the procedure to students, multiplying the bottom equation by four in order to set students up to eliminate y and solve for x . The teacher then walked through the procedure, first solving for x and then substituting the value she had found for x into one of the equations to solve for y .

Once the teacher finished solving for x and y , a student raised her hand and said she had done the problem differently. The teacher responded:

Teacher: You can, which way did you do it?

Student: I multiplied the top line by two.

Teacher: You multiplied the top one by two so you could get positive eight and negative eight. It doesn't matter which way. You can choose. Okay, you can choose which way you want to do it.

Here, the teacher explicitly reinforced the possibility of multiple pathways through the same procedure (in this case, multiplying the bottom equation by four as she had done or multiplying the top equation by two as the student suggested). Next the teacher paused and thought for a moment, and said:

Now, you could have, if you wanted to, you could have used the substitution method. You could have moved this $8x$ to the other side and y would have been equal to positive $8x$ plus 19. Then you would substitute it in. So you have options.

She continued on to show how students might have solved the same problem using a different method entirely—substitution—reinforcing that with a little algebraic manipulation this was a viable alternative and would yield the same solution. She next reminded students that they could also have considered using the graphing method, graphing each equation and finding their point of intersection. She cautioned students that the graphing method might not be the best method in all situations. She highlighted examples where, unlike the one she had presented, the solution did not contain integer values, cautioning students that in those cases it would be more difficult to “read” the solution off the graph than to derive it algebraically. In this segment of instruction, the teacher focused on not only teaching a particular procedure (the elimination method), but also in cueing students to the characteristics of the problem that motivated the selection of that method over others, while still emphasizing multiple other approaches to the same problem.

Observers also noted lessons in which this practice was completely absent. In these lessons, teachers presented specific procedures as absolute rules with little room for choice or thought to applicability. For example, in presenting a lesson on graphing linear equations, one teacher told students that they “*have* to first make sure your equation is in

standard form before you graph it.” While it was clear to observers that the goal of this particular lesson was graphing from equations in standard form, the idea that there were other procedures for graphing a line were not only omitted, but students were made to manipulate equations that could more easily be graphed in other ways (e.g. using point slope form). Similarly, in a lesson on solving systems of equations by elimination, one teacher told students repeatedly and explicitly that the coefficients must be the same and their signs must be opposite in order to be able to eliminate one variable. She never acknowledged the possibility of subtracting the two equations. This rigidity was accompanied by an absence of discussions or exploration of alternative solutions pathways.

Organization in the presentation of procedures. Even in the absence of making meaning around procedures, most researchers agree that instruction on procedures should be clear and well-organized. At a minimum, we would expect that when teachers present a procedure, the steps of the procedure are correct, explained clearly—either verbally or in writing—and organized in a way that allow students to replicate the procedure on their own. Observers noted that in some lessons in the sample, teachers presented procedures in organized and systematic ways, highlighting and clarifying mathematical information.

For example, in one lesson, a teacher presented the procedures for factoring trinomials and multiplying binomials. In one segment of instruction, the teacher reviewed the procedure for factoring $x^2 + 6x + 9$ using the box method. She began by asking, “what’s the first step I make?” A student responded that she should draw a box. The teacher then drew a two-by-two matrix on the board to the right of the problem (see Figure 2) and continued:

Teacher: Now I have my rectangle drawn. Can you help me out Student J? What should I put in that first box?

Student J: x^2

Teacher: Thank you. [Writes x^2 in top left quadrant of the box]. Now Student D, what goes into the next box? We have this equation...

Student D: the nine ...

Teacher: The nine. Exactly right. [Writes 9 in bottom right quadrant of the box].

Student J: Mister—three and three.

Teacher: Okay, Student J says I should put three and three here. Student J, how did you decide that?

Student J: Because three times three is nine...

Teacher: Three times three is nine and also, do I just put a three or do I ...

Student: $3x$

Teacher: $3x$ and $3x$. [Writes $3x$ and $3x$ in the top right and bottom left quadrants of the box]. Okay, so I've got my x^2 and my $3x$. Now what do I start doing?...

Student: For x^2 put one x on top and then one x on the side.

Teacher: [Writes x on the outside of the box above the x^2 and writes another x outside of the box to the left of the x^2]. Okay how'd you figure that out?

Student: Because x times x is x^2

Teacher: That's right. Because this is multiplying, x times x is x^2 . And what did you put on top of here? [Points outside the box, on top of the $3x$ in the top row].

Student: Plus three.

Teacher: Plus three. [Writes $+ 3$ above the $3x$ in the top row]. And what do we got down here? [Points to the $3x$ in the bottom row].

Student: Another plus three.

Teacher: Another plus three. [Writes $+ 3$ to the left of the $3x$ in the bottom row]. Alright. Now, is this my answer, just like this, this box?

Students: No

Teacher: No it's not. What form should I put this in?

Students: Expanded

Teacher: Expanded, or factored form. Very good. $(x+3)$ and $(x+3)$. [Writes $(x+3)(x+3)$ underneath the box].

In this example, the teacher talked the students through the steps of a procedure without attention to meaning or the concepts underlying the procedure. However, she was verbally clear about each step and used a visual aid, which she filled in sequentially and

logically, using gesture for emphasis at particular points in the procedure. Once she and the students determined the factors, she wrote the complete factored solution underneath the box. This presentation of the procedure clearly presented the information and allowed students to follow the steps, which they did in subsequent student work-time, repeating the procedure with little apparent difficulty.

In some cases, the presentation of procedures was exceptionally well-organized or systematic. For example, in one lesson on solving systems of linear equations using substitution, the teacher explained how to solve a system through a worked example. Then, she presented students with a new example. While she solved it with help from the class, she simultaneously wrote the steps to the right of each line of the worked algebraic example (See Figure 3). She told the class:

- Teacher:* First, identify the variable. It could be any variable, x or y , that has a coefficient of one. So the example would be like this— x or y . [Teacher writes Step 1 to the right of the problem]. A coefficient is the number in front of a variable. This is the coefficient in this term [points to $3x$]. It's three. In this term [points to $5x$] it's five. Here [points to x] it's one. So you look for the variable that has – like we say in English—no number. It really does have a one but there's no number next to it. That's the easy way. So once you have identified it and we know it's x . Right? You take this [circles $2y - 4$] and what do you do with it?
- Student:* You replace it.
- Teacher:* Where?
- Student:* You replace it with the x .
- Teacher:* [Points at x]. Yes. Say it again.
- Student:* Replace it with the x .
- Student:* Ohhhhhhh.
- Teacher:* Yes. Three, open [parentheses] $2y - 4$. you see? I replaced the x with the $[2y - 4]$.
- Student:* Why?
- Teacher:* Because now, everything is in y terms. You see? By doing it in y , now I can solve. Before if I only had this equation—like this one [points to $3x - 5y = 11$] I can't solve it because I don't know what x or y is. If I have one variable, then I can solve. Here they're all y 's so then I can solve.

- Student:* So what's the next step?
- Teacher:* So what's the next step?
- Student:* Now we solve.
- Teacher :* So you find this one, you replace it. Replace the variable and solve. [Writes Step 2 to the right of the problem]. Okay, so now we solve. What do we get?
- Student:* So we always solve for the one that has no coefficient.
- Teacher:* Yes! But that's incorrect. Even though it looks like it has no coefficient it really is one.
- Student:* Alright but—
- Teacher:* But that's what you were meaning right?
- Student:* Yeah.
- Teacher:* Okay, so in like regular English language, you look for the one that has no number in front of it, but really in math language you have to be very careful. This is a one coefficient. Okay? Okay.

The teacher went on to write out the expression in terms of y and solve for y , finding an answer of $y = 23$. In demonstrating the procedure, she explained each step clearly, answering student questions about how the substitution worked and underscoring mathematical language. She used the written record, along with circling and arrows to draw students' attention to the steps of the procedure. After solving for y , the teacher reviewed the first two steps in the procedure she had written on the board and continued the problem:

- Teacher:* We came in and we substituted the x . Now I only have y 's in my equation. And then by simplifying and solving for y , I get $y = 23$.
- Student:* So all you did was solve for y ?
- Teacher:* Yeah. Now how do we find the x ?
- Student:* Plug in the y
- Teacher:* Yes... The last step. Replace your number, your value in the original equation. [Writes Step 3 on the board].
- Student:* Oh—these steps make this so much easier.
- Teacher:* It could be any of these equations—either of the original two. Which one do you want?
- Student:* The bottom one.
- Student:* When you write out the steps I get it.
- Teacher:* Yes. Everybody does.

In this case, the teacher not only laid her procedural thinking out clearly and explicitly, she used the written record to reinforce mathematical language and some of the concepts inherent in solving with substitution. Students responded positively to this level of organization (“Oh—these steps make this so much easier” and “when you write out the steps I get it”).

Not all lessons included such high levels of organization however. In some lessons, the presentation of procedures was relatively disorganized, seeming to result in student confusion. For example, in one lesson on graphing linear equations, the teacher told the students to graph $y = 4 - 3x$. She wrote nothing on the board and quickly verbalized the procedure, saying, “We solve for y in terms of x if it’s not already in that form... Then we choose at least three values for x . Do four or five... You want to make sure you have a straight line... Pick zero and one for sure as two of your x -values. Compute corresponding values for y , plot the ordered pairs, draw the line through them.” Students appeared confused and one student said, “I thought it could be any number?” The teacher replied, “It can be, you can put anything you want in for x , but I’m telling you to use zero and one as two of your x ’s. Create your table, plot your points and graph it.” The teacher’s rapid fire and at times contradictory verbal directions (choose three values for x , followed by choose four or five), accompanied by nothing written on the board, appeared to confuse students who then struggled for over eight minutes to graph the line.

Leveraging connections to build understanding. Students frequently struggle to see how their prior knowledge relates to emergent algebraic understanding (Booth, 1988). In addition, the abstract nature of algebra—its focus on symbols, form, and structure—

may present further obstacles to student understanding (Hiebert & Grouws, 2007; Rakes, et al., 2010). Finally, students may see algebra as a series of abstract and disconnected topics, lacking coherence (Thompson, 2013). Irrespective of whether teachers are focusing on concepts or procedures, there is promise in making explicit connections—both within and outside algebra content—between representations, objects, and topics.

In this sample, observers noted the ways in which teachers referenced or made strong use of connections in teaching algebra, regardless of whether they were teaching procedures, concepts, definitions or other aspects of the curriculum. These connections may be important because they can provide mathematical meaning to algebraic abstraction and can potentially attend to student difficulties with algebra. Three specific types of connections emerged from the analysis. First, teachers made connections across representational forms, specifically graphs, tables, equations, and contexts. Next, teachers connected the algebra in the lesson to other aspects of the algebra curriculum or the broader domain of mathematics, situating the mathematics under study. Finally, teachers made connections between abstract algebraic ideas and concrete content or examples, connecting newer algebraic ideas to more familiar concrete properties, examples, or mathematical objects.

Connections across representations. The emphasis of the mathematics reform movement on both a functional view of algebra (Kieran, 2007) and the accompanying use of multiple representations (Knuth, 2000) was apparent in algebra classrooms in this sample. Teachers frequently presented multiple representational forms where appropriate (e.g. graph, table, equation, and problem context). In some lessons, these representations were simply presented verbally, in the text, or in writing, without any attempt to connect

them. However in many lessons, teachers at least referenced the connection between the representations, and in some cases explicitly reinforced important connections.

For example, in a lesson on factoring binomials, after factoring an expression, the teacher remarked to students that if she were to enter the expanded and factored forms of the equation into a graphing calculator and graph the two equations, she would see the same parabolic graph. This statement, although brief, not only reinforced the equivalence of the standard and factored forms of the equation, but in the context of a manipulation-heavy lesson, also reminded students of the connection between the equation and its graphical representation. In a lesson on graphing linear equations, another teacher reinforced the relationship between the table, graph and equation. She explained to students that the numbers in the table generated from an equation provided the coordinates for the graph of the line and were also the solution set to the algebraic equation. She emphasized the connection, pointing between the graph and the table, stating, “every point on this line has a corresponding x and y coordinate.” She next pointed at the equation and said, “Any point on this line, the corresponding x and y coordinate will fit into that equation and make it work.” Finally, she gestured between the table of values, the equation, and the graph as he said, “This is the solution set to that linear equation. A linear equation graphs out as a ... straight line.” Again, this type of connection worked to both give mathematical meaning to linear equations and functions and to help students navigate a potential point of confusion around the form and structure of linear equations.

In addition to brief connections such as the preceding example, there were also instances in which teachers explored the connections between representations in depth or

with great detail. In one lesson on transformations of parent graphs, the teacher presented the equation of an absolute value function and the graph of $y = |x|$ and talked with students about the ways in which the equation would change if the graph shifted in various ways (up three units, to the right two units, etc.). She then discussed the ways in which the graph would move if the equation were written with different constant values (e.g. $y = |x + 3| - 2$ or $y = |x - 3| + 1$). While the content of this lesson may have lent itself to more substantive connections across representations, the teacher did significant work to relate the behavior of the graph to the values in the equation.

Topics that might naturally exploit connections across representations did not always feature instruction that did so, however. For example, many lessons in the sample focused on solving systems of equations, a topic that might lend itself well to connections between representations, particularly since one canonical method for solving systems utilizes graphing while other canonical methods are algebraic. In some lessons on solving systems of equations, the algebraic procedures for solving the system (e.g. substitution or elimination) were presented as distinct from the graphical method for solving the system. In other cases, the graphical and algebraic methods were compared briefly, with teachers mentioning, “you would get the same answer if you graphed these.”

Situating the mathematics. Another type of connection featured in the sample was connections across topics in the mathematics curriculum or in the broader domain of mathematics. These connections have the potential to help students see how what they are learning fits into a larger context and motivate the topic under study, particularly when it is relatively abstract (Thompson, 2013). We saw evidence of teachers situating and motivating the mathematics in a given lesson in the context of prior mathematical

knowledge, connecting a particular topic to future content, making connections that developed a mathematical through line within and among lessons, and connecting current content to the larger domain of mathematics.

One way in which teachers situated the mathematics was to make active connections to prior content, linking the mathematics the class was currently learning to other topics the class had studied. For example, in a lesson in which the teacher was introducing the procedure for solving systems of linear equations using elimination, she paused after eliminating one variable and remarked to students, “Now that we have eliminated one variable, this equation looks like the single-variable algebra equations we already know how to solve.” In another lesson on graphing quadratics, after the students located the roots of the parabola on the graph, the teacher took the opportunity to connect the idea of roots to the x-intercept in a linear equation—both graphically and through the algebraic process of setting $y = 0$ —noting similarities and differences between the new material and what students already knew.

Teachers also grounded new content by making connections to future topics of study, previewing how the day’s lesson connected to later topics in the curriculum. For example, after introducing how to plot points from an equation onto the coordinate plane to generate a line graph, a teacher had the class complete a number of examples by themselves. As she was circulating the room, she commented to the class how this exercise set them up for future units of study, particularly the connection between graphing points and graphing lines. She acknowledged the challenges students were having with graphing and said, “After this topic, we’re going to go into slope. So what you’re learning right now is the beginning of slope. Same thing. You have an equation,

you give values, you graph it. Then we will talk about whether it is a positive slope, negative slope... That's the next topic we're going to do, this is the introduction to slope."

Another way in which teachers situated the mathematics was by creating mathematical through lines within lessons, explicitly linking the sections of the lesson together and highlighting the connections between the component parts. For example, one teacher began a lesson on systems of inequalities with a warm-up problem that asked students to find the slope of a line between two given points and then write the equation for the line that passed through those points. The teacher then explicitly articulated the mathematical purpose of the warm-up in preparing for graphing systems of linear inequalities. The teacher built upon the solution to the warm-up, making clear connections between how understanding that material allowed students to graph and solve systems of equations. The teacher then made another explicit connection between how the processes for solving systems of equations would allow students to be able to solve systems of linear inequalities. Here, the teacher focused not only on activating prior knowledge (graphing linear equations) and connecting that knowledge to a new topic (solving systems of linear inequalities), but also generated a coherent mathematical through-line within and among a series of lessons. This type of connection may work to help students to better understand the difficult shifts within the algebra curriculum.

Teachers also situated the mathematics by making connections to other mathematical topics or to the broader domain of mathematics. For example, in one lesson, while displaying how changing the values of m and b in a linear equation of the form $y = mx + b$ changed the shape of the graph, the teacher commented that such transformations

will reappear in the context of higher degree functions later in students' mathematics careers.

While situating the mathematics occurred to some degree across many lessons, there were some lessons in the sample in which material was presented as discrete, compartmentalized, and disconnected from the broader domain of algebra. For example, in a lesson on trigonometric ratios, the teacher began with a warm-up in which she asked students to answer some identification questions about a right triangle with the side lengths labeled and some side measures included (e.g. How long is the side opposite angle R? Which side is adjacent to angle A? What is the length of the side opposite angle R divided by the length of the hypotenuse?). After recording students' answers to the final question, the teacher immediately transitioned to presenting the three trigonometric ratios, writing sine, cosine, and tangent on an overhead projector transparency. She did this with only a passing reference to the warm-up, but made no connection to either how what the students had done earlier in the lesson reflected a trigonometric ratio or to the role of the right triangle in generating these ratios. She wrote on the overhead that $\sin\theta = \text{opposite/hypotenuse}$ and said:

What we did is find the sine of an angle. We abbreviate this s-i-n. We find the sine and we always have an angle here. I'm going to use this funny symbol for an angle. It's called theta. It's a Greek letter and that just means it's an angle measure. So the sine of the angle measure is equal to the length of the side opposite the angle over the hypotenuse.

She continued in this manner, defining the three trigonometric ratios by their formulas and giving students problems with missing side lengths to solve.

Students used the trigonometric ratios to find the missing side lengths, but were

given no opportunity to see how this new topic fit into the larger mathematical storyline.

Making connections between concrete and abstract ideas. One of the main reasons students struggle with algebra is that it is more abstract than the mathematics they have previously encountered (Booth, 1988; Hiebert & Grouws, 2007; Rakes, et al., 2010). If teachers can assist students in making connections between algebraic abstractions and their numerical foundations using concrete examples, representations, or ideas, students may be better able to combat some of these difficulties and build their understanding (Kieran, 2007). We saw some evidence of this practice across lessons, particularly in the ways in which teachers leveraged concrete examples, representations, or ideas to develop understanding of abstract concepts, formulas, notation, and definitions.

One frequent way in which teachers attended to abstraction was in underscoring what abstract symbolic notation represented in algebraic formulas, unpacking the components of a formula by attending to what the symbols in the formula represented. This was frequently done with labeling such as in a lesson on the properties of exponents in which one teacher explicitly identified the base and exponent in each property, labeling each while discussing their roles in the formula. In a more sustained example, one teacher developed the formula for arithmetic sequences, engaging in lengthy and careful work to have students identify the initial value and common difference in each example problem she gave. She built to an algebraic expression for calculating any term in a sequence and made concrete each term in the formula by connecting to the initial value and common difference in the earlier examples.

Another way in which teachers helped students make sense of abstract ideas was to clarify mathematical definitions or abstract algebraic concepts using concrete examples. For example, in a lesson on angle relationships and lines, after defining each relationship (e.g. supplementary angles, parallel lines), the teacher had students develop their own examples and non-examples of each definition. This activity asked students to interpret the definition through a concrete example of their own making. In another lesson in which students simplified rational expressions using the properties of exponents, the teacher went over a problem that asked students to simplify $\frac{x \cdot x \cdot x \cdot x \cdot y \cdot y \cdot y \cdot y}{x \cdot x \cdot x \cdot y \cdot y}$. To motivate the process for doing so, she asked students what $5/5$ and $2/2$ simplified to, explaining that just as a number divided by itself equaled one, an unknown divided by itself also equaled 1. She leveraged this understanding to explain the process of “cancelling out” the appropriate number of x ’s and y ’s in the process of simplification, reiterating for students, “so when I cancel these out, they don’t turn into a zero, it’s one” as the reason that the simplified form of the rational expression is x^2y^2 . Less frequently, teachers did this process in reverse, at times leveraging abstract algebraic tools or representations to solve concrete examples. In one lesson, the teacher presented a problem in which students had to find the distance between two buildings on a city map. The teacher superimposed a right triangle onto the map and used this to reinforce the role of the Pythagorean Theorem to find the distance.

Teachers also attended to abstraction by using analogies to connect to concrete ideas or topics that students had previously mastered. For example, in a lesson on solving systems of equations using substitution, the teacher presented the students with two equations: $y = 2x - 3$ and $2x - y = 5$. The teacher then asked the students if they would be

willing to trade one dollar for four quarters. When the students said they would, the teacher asked them why. The students replied that they were the same amount of money so it did not matter if they had the four quarters or one dollar. The teacher used their response to motivate the idea of substitution, saying that it was possible to substitute the y in the second equation with the $2x - 3$ from the first equation because they had the same value. She told the students that the $2x - 3$ was “the four quarters” and the y was “the dollar,” showing how substituting one for the other did not change the value. Teachers also leveraged mathematical analogies. For example, in a lesson on factoring, the teacher connected the process for finding factors of a polynomial to factoring a numerical product, a concrete mathematical idea for students.

In some instances, this practice capitalized on one abstraction with which students were familiar to make sense of a new one. For example, in one lesson steeped in algebraic manipulation, students were struggling to utilize the rules for simplifying rational expressions that included radicals. The teacher helped students to make sense of the unfamiliar rules by relating them to how variables operate, a concept with which they had a good deal of familiarity. She reminded students that they knew that $x + x$ equaled $2x$. Similarly, $\sqrt{3} + \sqrt{3} = 2\sqrt{3}$. She reiterated that just as $x + y$ could not be combined because they are unlike terms, students should not combine $\sqrt{3} + \sqrt{2}$. However, she made the connection that just as x times y can be expressed as xy , so too can $\sqrt{3} \cdot \sqrt{2}$ be expressed as $\sqrt{3 \cdot 2}$ or $\sqrt{6}$.

Finally, we also saw teachers leverage concrete examples by using numbers to develop general algebraic rules and properties. For example, in a lesson on the properties of exponents, one teacher worked with students to develop the rule for solving problems

with negative exponents. She first asked students to solve a series of numerical exponentiation problems (3^4 , 3^3 , 3^2 , and 3^1) and then extended the pattern for (3^0 , 3^{-1} , and 3^{-2}). She used this pattern to help students develop intuition around the rule $x^{-n} = \frac{1}{x^n}$.

While this does not attend to the meaning of why a base raised to a negative exponent is equal to its reciprocal raised to the positive exponent, it does ground the seemingly abstract rule in the concrete number system with which students are familiar.

Other Practices Prevalent in the Literature

While instruction on procedures and leveraging connections as described above were the most salient features in the data, it is worth noting that other instructional features prevalent in the mathematics literature also appeared in the sample, although these practices occurred sporadically and often were enacted without much depth. The most common of these practices was the mention or use of contextualized problems or real-world situations. Five of the 42 lessons in the sample featured in depth work around contextualized problems for part of the lesson (it is important to note that two of these were the lessons coded as inquiry/exploration focused). In an additional four lessons, teachers noted real-life examples of the mathematics being presented, though contextualized problems were not worked on or solved. For example, when telling students why they needed to learn how to solve a system of equations, one teacher referenced comparing cell phone plans for the best value.

Other reform-oriented practices appeared less frequently. For example, four lessons featured the development of a generalization or formula from a pattern (e.g. using a number of iterations of an arithmetic sequence to develop a formula for finding the n th term in the sequence). Less frequent, but notable in the sample, were lessons that

included a significant amount of time focused around making sense of a concept (e.g. the difference between a function and a relation) as distinct from doing so in the context of learning a procedure (three lessons), and lessons in which students were given non-routine and cognitively challenging tasks (two lessons).

Discussion and Implications

In this exploratory study, I conducted a grounded analysis of algebra lessons to identify and describe the format and key features of instruction in contemporary ninth grade algebra lessons from five urban districts. Similar to other large- and small-scale studies of mathematics classrooms (e.g. Hiebert, 2013; Hiebert, et al., 2005; Star, et al., 2000), I find that the predominant mode of instruction is teacher-centered with some limited opportunity for student practice and participation. Of the 42 lessons in this sample, 34 lessons followed this instructional format. Furthermore, the teaching of procedures was frequent in this sample with 38 of the 42 lessons featuring at least some time in which students were taught procedures. Moving beyond the format of the lesson, I also identified particular instructional features that were prevalent in the sample, describing these practices in light of how they might theoretically work to improve student understanding of algebra content. Specifically, features along two different categories emerged—instruction on procedures and leveraging connections to build understanding. I identified three instructional features prevalent when teaching procedures: the ways in which teachers and students made sense of procedures, the ways in which teachers supported students in developing flexibility in their use and choices of procedures, and the ways in which teachers organized their presentation to make clear the steps of procedures. I also identified three practices by which teachers make connections in the

context of their instruction that may work to both promote student understanding and potentially counteract student difficulties with algebra. First, teachers made connections across representational forms. Second, teachers situated the mathematics by connecting it to prior or future knowledge, other topics in algebra or the larger domain of mathematics. Finally, teachers made connections between abstract algebraic ideas and concrete objects and representations. I described what these features looked like in practice drawing on examples from lessons from the sample, and also described lessons in which these features were notably absent.

While the features that emerged from this analysis were theoretically grounded and germane to algebra instruction in particular, I also acknowledge that they are not the only features of instruction common to algebra lessons. It is in fact quite likely, if not certain, that there are practices important to student learning and achievement in algebra not captured in this analysis. For example, general classroom practices (e.g. time on task, classroom management, or classroom climate) and more general mathematical practices (e.g. teachers' mathematical explanations or how teachers remediate students' mathematical misunderstanding) may be equally important and could certainly have been a viable theoretical lens through which to analyze this data.

The findings of this analysis indicate the importance of looking closely at classroom instruction. In discussing contemporary algebra classes, Hiebert (2013) remarks, "your grandfather would likely recognize the math class your children attend" (p. 45). While the results of this analysis show further evidence of the persistence of more traditional pedagogical formats where instruction remains heavily teacher-centered and focused on procedures, stopping the analysis here would both obscure important changes

that have begun to take root and hinder the development of potentially promising pedagogical practices. I am not advocating for abandoning the push for ambitious instructional practices. Rather, I argue that *in addition*, there may be promising practices that teachers are already engaging in that can be leveraged for instructional improvement.

Many scholars have commented that the format of teaching may be deeply culturally engrained (Cohen, 2011; Cuban, 2013; Hiebert, 2013) and difficult to change as it is learned through observation and handed down from teacher to student (Hiebert, 2013; Lortie, 1975). Indeed, there is evidence across disciplines that many attempted instructional reforms are simply incorporated into teachers' existing schema, resulting in technical elements being grafted onto established pedagogical routines (Cohen, 1990; Cuban, 2013). In mathematics, for example, instructional enhancements such as the use of manipulatives can be used in either a teacher-directed or student-centered mode of instruction and be used in the service of developing either procedural or conceptual knowledge (Rakes, et al., 2010). Thus a focus solely on format may obscure potentially beneficial practices that can be instantiated regardless of instructional method.

In this vein, rather than focus improvement efforts on what appears to amount to a sea change in instructional format, Hiebert (2013) advocates for a deeper examination into teaching *practices*, turning teaching itself into the unit of analysis and engaging in careful investigation into its component parts. By engaging in close analysis of videos of algebra lessons and describing features of instruction in this sample, I find more nuance than a focus on instructional format might indicate. For example, teachers in the sample engaged in practices that afforded students the opportunity to make meaning of procedures, and develop procedural flexibility and efficiency. They also emphasized

important connections across representations, topics, and ideas that have the potential to counteract many of the theoretical difficulties students have with algebra. While these practices occurred at varying levels of quality across lessons in this sub-sample, that they were salient is cause for hope. Thus the results of this study provide evidence of potentially promising practices that *teachers are already engaging in* that can be further leveraged for improving student understanding in algebra.

This optimism however requires a theoretical shift away from seeing teaching procedures as somehow opposed to or unrelated to teaching concepts. Indeed, flexibility and efficiency with procedures may be an important aspect of algebra in particular, as the subject contains a significant amount of procedures of symbolic manipulation (Kieran, 2013). Procedures are not only prevalent in algebra, but they are often more complex than those in arithmetic (Star, 2005; Star & Rittle-Johnson, 2009). In fact, deep procedural knowledge—using procedures flexibly and efficiently—may be highly beneficial to student understanding in algebra (Star, 2005). When thinking about algebra, it may in fact be more useful to think about its symbolic procedural activities *in terms of concepts*. In algebra, procedures may be conceptual in nature—both in their initial introduction and as they are revised to encompass new mathematical ideas and situations (Kieran, 2013).

Like Kieran (2013) and Star (2005), I see the separation of procedural skills and conceptual understanding as somewhat problematic, and especially so within the context of algebra. First, teaching procedures is so prevalent in algebra that it warrants a deeper investigation and second, instruction on procedures may be necessary for developing skills, knowledge, *and* understanding in algebra. This separation also encourages the equating of teaching concepts with “good” teaching, which by extension implies the view

that teaching procedures is somehow “less good.” While many mathematicians and mathematics educators would not claim to argue this stance, it is one that appears periodically, most recently in the popular debate surrounding the Common Core State Standards (see for example, Strauss, 2014). While Star (2005) argues that both concepts and procedures can be taught superficially or in depth, I see them as less easy to separate. In this analysis, I showed examples of teachers making meaning of procedures such that students had the opportunity to develop both facility with a given procedure and an understanding of the concepts underlying the procedure. By articulating and unpacking specific features of teaching procedures, we may be better positioned to understand and describe what this instruction looks like at high quality also to assist in building teachers’ capacity around instruction on procedures.

Hiebert and Grouws (2007) suggest that explicit and public connections between and among mathematical facts, procedures, and ideas promote conceptual understanding. This analysis shows specific ways in which teachers promote these connections in an algebra classroom. For example, in making connections across representations, teachers afforded students the opportunity to better understand the relationships between different structural forms. In situating the mathematics, teachers frequently articulated the goal or purpose of the algebraic topic under study and made clear to students how it fit in a broader sequence of topics or how it connected to other mathematical ideas. Doing so may theoretically work to motivate the topic under study and ground new content in a mathematical through-line. Finally, teachers in the sample made explicit connections between abstract algebraic ideas and notation to more concrete representations and objects, facilitating the cognitive shift into unfamiliar form and structure.

It is important to note that there are also theoretically promising instructional practices identified in the literature that did not appear in the data. For example, Hiebert and Grouws (2007) advocate that teachers allow students to struggle with important mathematics and actively grapple with key ideas. They and others advocate for tasks high in cognitive demand as well as significant opportunities for student discourse and analysis. This exploratory analysis did not uncover evidence of these practices and there is reason to believe they have not taken root at scale. Using a number of established classroom observation instruments, the larger MET study found minimal evidence of cognitively challenging tasks, active student engagement, or deep analysis in mathematics classrooms (Kane & Staiger, 2012). Instead, both this and other analyses of MET data (see for example, Kane & Staiger, 2012; Litke, 2014) revealed a good deal of student passivity with brief periods of engagement in doing mathematics, most often in the context of practice problems that replicated teacher-presented examples. More research is needed to see if these practices are relatively absent in larger, more representative samples and if so, why they seem difficult to instantiate.

This study relies on data from five urban districts and as such, results may not generalize to other districts or regions. This may not be terribly problematic as the goal of this study was to articulate extant practices that teachers employ rather than make claims about the prevalence or quality of those practices. In addition, the teachers who participated in the MET study do not constitute a representative sample, as districts and teachers volunteered to participate. Thus, observed instruction may not be representative of all teachers and within teachers, video-recorded lessons may not be representative of regular practice. Again, as I am interested in exploring teaching rather than making

generalized claims about teachers, this may not be that troubling of an issue. Furthermore, given the ubiquity of instruction on procedures in algebra and the push in the mathematics education literature for the importance of connections, the results of this analysis are instructive for both efforts to support teachers and for motivating further research. It is also important to acknowledge that I purposefully selected a stratified random sample from among the ninth grade algebra teachers based on CKT score quintile, in order to maximize potential variability of practices. Thus the frequency and quality of the instructional features in this particular sample may not be an accurate reflection of how they might appear in this data set. Further research on more representative samples is certainly needed to test whether the practices identified in this study are germane to more teachers in other settings and to determine the level of quality with which they are enacted. That teachers are utilizing these practices even superficially is promising and suggests a potential path for instructional improvement. If we can articulate what these features would look like at high levels of quality, it may be possible for teacher educators to support teachers in incorporating them where appropriate into their instruction.

The identification of instructional features is an important first step in supporting quality algebra teaching. While these particular features may theoretically support student learning in algebra, we do not know yet whether that is the case in practice. As of yet, there is no indication as to whether there is a relationship between these particular practices and student outcomes. The findings of this study can be used to inform future research regarding questions of improving student understanding in algebra. For example, one possible direction would be to develop a framework for measuring the frequency and depth of these practices in classrooms and to use this framework to investigate their

impact on student understanding and achievement. Another avenue for further research is to determine whether these particular features are salient in larger or more representative samples.

Conclusion

The MET study provides one of the largest snapshots of classroom teaching we have to date and presents an extraordinary opportunity to explore the qualities of teaching across classrooms. While teachers are increasingly held accountable for their students' performance and growth in mathematics (Hill & Grossman, 2013), their instruction goes comparatively unexamined. This study looks inside the algebra classroom to focus on algebra *teaching*. Given decades of reform and the ambitions of the Common Core, the paucity of research exploring contemporary instruction is particularly striking. If our understanding of what it looks like to teach algebra is underdeveloped, then we must do more to advance our knowledge of actual features of instruction in order both to understand what practices might be most effective and to provide teachers with support to develop these practices. With this study, I hope to broaden the discussion of school algebra beyond a framework of idealized practices to one that considers the nature of the instruction students actually experience. Better understanding contemporary instructional practice furthermore allows for a clearer sense of the distance between where we are and where we want to be, particularly as states and districts embark on implementation of the Common Core State Standards. In addition, a better understanding of extant instructional practices provides insight for those hoping to improve instruction, providing a potential starting point for efforts at improving teaching quality in algebra.

Chapter 3.

Measuring Instructional Practice in Algebra: The Development of An Observational Instrument

In the previous chapter, I identified instructional formats and features that were salient in the sample of algebra video. The results of this exploratory analysis revealed specific practices that teachers engaged in that had the potential to be leveraged for improving student understanding in algebra. Yet it remained an open question whether and to what degree of quality these practices are prevalent in a wider sample. In order to answer this question, a systematic protocol for examining and describing instructional practice along these particular dimensions is necessary. Thus, using the results of the analysis in Chapter 2, I developed an observational instrument, the Quality of Instructional Practice in Algebra, keyed to the instructional formats and features elaborated therein.

In this chapter, I have three broad goals. First, I make the case for the creation of an algebra-specific instrument, looking closely at how teaching quality has been measured in the literature, both in general and in mathematics specifically. Next, I discuss the instrument itself, describing the instrument-development process, presenting a description of the codes used in the instrument, and describing the scoring procedures with examples from classroom practice. Third, I discuss the quality of the information yielded by the instrument, presenting information on the validity and reliability of scores from a sample of video coded with the instrument. I conclude with a discussion of the utility and potential of such an instrument for both future research and for supporting teachers through professional development.

Observing and Measuring Teaching Quality in Algebra

In order to better understand the nature and quality of instruction in algebra lessons broadly, a classroom observation instrument specific to algebra is likely useful. Classroom observation instruments allow trained raters to document particular dimensions of instructional practice and measure their level of quality against pre-determined benchmarks. Researchers advocate for the use of such observational instruments to both evaluate and improve instruction (Kane & Staiger, 2012; Pianta & Hamre, 2009), while others rely on standardized protocols to describe instructional practice (e.g. Hiebert, et al., 2005; Hill, et al., 2014). Successful observation systems can create specific information about the quality of teaching practices with the goal of improving teaching so as to improve student learning (Gitomer, Bell, Qi, McCaffrey, Hamre, & Pianta, 2014).

Observation protocols have evolved in both their aims and focus. Early observation protocols measured general dimensions and practices such as the quantity and pacing of instruction (including features such as classroom management, amount of content covered, and time on task), the format of teaching employed (e.g. direct instruction, whole-class vs. small group time), the nature of teacher talk and questioning, and classroom organization (Brophy & Good, 1986). These systems were largely agnostic to content, measuring little about the mathematics-specific teaching in lessons (LMTP, 2011). More recently, scholars and instrument developers have argued that understanding and assessing the quality of teaching requires attention to domain and subject-specific aspects of instruction (Grossman, Loeb, Cohen, & Wykoff, 2013; Stodolsky, 1988). As a result, content specific protocols have been developed in

mathematics (e.g. MQI) and English Language Arts (e.g. PLATO), as well as those that have developed to cover multiple content areas (e.g. RTOP) or with discipline-specific foci (e.g. UTOP).

Existing mathematics-oriented protocols come from the perspective that the mathematical work that occurs in classrooms is distinct from classroom climate, pedagogical style, or the deployment of generic instructional strategies (Hill & Grossman, 2013). Yet even within this framing, existing protocols must necessarily focus on particular practices and perspectives within the discipline of mathematics. For example, the Mathematical Quality of Instruction (MQI) orients its protocol to assess the quality and richness of the mathematical content offered to students, the ways in which teachers and students interact in the context of the mathematics, and the degree to which students engage meaningfully with the mathematics in a given lesson (LMTP, 2011). The UTeach Observation Protocol (UTOP) focuses somewhat differently on broader elements such as classroom environment, lesson structure, and lesson implementation, as well as content-oriented elements such as accuracy, the appropriate use of symbols and representations, relevance of the content, and the ways in which the content of the lesson relates to other discipline-specific content (Marder, et al., 2010). Also of note is that existing mathematics-oriented protocols are either designed to capture teaching quality in the elementary and middle grades (e.g. MQI; TIMSS video protocol) or are intended for use across grade levels (e.g. UTOP) and STEM content domains (e.g., Reformed Teaching Observation Protocol, which captures both mathematics and science). As such, they may not capture important dimensions of instructional practice specific to algebra. While

some of the specific practices identified in Chapter 2 are embedded in particular codes in the above instruments, they are often included in the context of larger categories.

Some instructional features identified in the exploratory analysis presented in Chapter 2 are largely absent from existing instruments. For example, I find that instruction on procedures is prevalent in algebra, and identified particular instructional features that may support the development of students' algebraic understanding (such as supporting procedural flexibility and making sense of procedures). Yet many existing instruments are designed to detect reform-oriented instructional practices, focusing for example on student exploration, the nature and proportion of student talk, connections with real-world phenomena, or promoting conceptual understanding (see for example, Piburn, et al., 2000) and thus may not be attuned to differences in quality with respect to instruction on procedures. In fact, many such instruments treat procedural instruction as a deficit (for example, as the default mode of instruction when conceptual instruction is not present) and thus are not designed to attend to the nuance of how and how well procedures are taught. Other dimensions of practice particularly salient to algebra instruction may not be adequately captured by existing instruments, including connecting prior mathematical understanding to new learning (Booth, 1988), and connecting abstract ideas to concrete underpinnings. These practices may be important to attend to in algebra instruction as algebra represents a shift toward mathematical abstractions and symbolic representations that can present challenges for students (Star et al., 2015).

Other reasons to develop an algebra-specific instrument connect to potential future uses of this instrument. The first such reason relates to the increasing heterogeneity of skill levels among those enrolled in the course. As access to algebra has increased,

there remain concerns about gaps in both student achievement and the quality of instruction (National Mathematics Advisory Panel, 2008). There is evidence that the quality and qualities of instruction in classrooms may vary by student skill level and demographic composition (e.g. Diamond, 2007; Gutstein, 2006; Litke, 2014; Smith, Lee, & Newmann, 2001), for example with students from disadvantaged backgrounds provided *less* access to reform-oriented mathematics instruction than their more privileged peers (Payne, 2008, Smith, et al., 2001) and higher achieving students more likely to encounter reform-oriented practices (Litke, 2014). Identifying algebra-specific features of instruction and determining indicators for various levels of quality allows for future investigations into whether these particular features are equitably distributed across students.

Second, using an algebra-focused observational instrument to describe the nature and quality of specific features of instruction has implications for instructional improvement. Traditionally, such instruments have been used to evaluate teaching quality (e.g. Hill, et. al., 2008) and in some cases teacher quality (see for example Gargani & Strong, 2014 and Kane & Staiger, 2012). An instrument focused on more fine-grained instructional practices may have more potential for use with teachers for the purposes of instructional improvement, as it more specifically articulates what specific instructional features look like across levels of quality.

Finally observational instruments have been put to use as a systematic means to describe instruction (e.g. Hiebert, et al., 2005; Hill, et al., 2014; Jacobs, et al. 2006). The descriptive results generated by the TIMSS studies for example have been important to researchers' and policy-makers' understanding of middle school mathematics instruction

and have provided great insight into instructional improvement efforts. Studies describing preschool classroom environments using the Classroom Assessment Scoring System (CLASS) have likewise influenced thinking about early childhood education practices (e.g. Justice, Mashburn, Hamre, & Pianta, 2008; Mashburn, et al., 2008). Using an observation protocol to better describe contemporary algebra instruction can similarly inform improvement efforts.

It is important to note that an observational instrument focused on fine-grained instructional features in algebra will do little to aid in the evaluation of teachers, as it—by design—may have too narrow a focus and may not prove practical for districts (or even math departments) to develop and administer. Rather, the specificity of the instructional features captured by such an instrument may be more useful for providing targeted information to policy-makers or administrators, helping both to better describe and understand the variation in quality of instruction and provide guidance on where to focus resources and support for teachers. For example, a standardized instrument focused on describing instructional features and developed as a tool for instructional improvement may hold promise for professional development. Pianta and Hamre (2009) argue that direct assessments of teaching using standardized and validated instruments provides important information not only on the quality of the instruction students receive but also insight into the “science” of teaching. Similarly, such an instrument can also be used to evaluate the effectiveness of professional development activities, measuring whether and to what degree teaching has changed along these dimensions.

If such a protocol is to be of use in the above ways, specific embedded assumptions that must be satisfied that allow for the inferences intended. For example,

protocols used for instructional improvement must be able to meaningfully distinguish between different instructional practices in such ways that the judgments formed based on scores are useful. In addition, observers need to understand the framework in similar ways so as to make the same judgments about teaching quality from the same pieces of evidence (Bell et al., 2012; Gitomer, Bell, et al., 2014). Given the growing prevalence of classroom observation protocols, researchers have rightly called for investigations of the quality of the information yielded by these instruments (Bell, et al., 2012; Gitomer, Bell, et al., 2014; Hill et al., 2012). Bell and colleagues (2012) argue that attending to the validity of the scores produced by observational instruments is particularly important given the presumption of their uses for either the purposes of evaluation or feedback to teachers. While the purpose of this algebra-specific instrument is not evaluative, it is intended to provide information about the quality of instruction and thus similar questions about the quality of that information remain.

This paper thus seeks to further develop our understanding of the nature and quality of algebra instruction by describing the development of a systematic observational instrument and exploring the validity and reliability of scores generated by the protocol.

Method

In this section, I briefly describe the data sources used in the development and validation of the instrument. I next describe the sample used in the analysis. Finally, I describe the process by which I developed the algebra-specific observational instrument.

Data Sources

Classroom videos. Consistent with others who have engaged in instrument development (e.g. LMT, 2011), I relied on video-recorded classroom lessons. Video allows for a deep analysis of the features of classroom instruction, enabling researchers to slow down and re-watch classroom moments that may be missed in live observation (Hamre, et al., 2007; LMTP, 2011). Video-recorded lessons also allow for iterative analytic processes in which themes that emerge from the video are refined and applied to new lessons (Jacobs, et al., 1999). Consistent with this recommendation, I secured access to the video-recorded ninth grade algebra lessons from the MET study. Lessons were video-recorded using an un-manned panoramic digital video camera operated remotely by teachers or school personnel (Bill & Melinda Gates Foundation, 2010a). While teachers selected the lessons to be recorded, the project requested that at least two of the lessons be focused on core topics in the teacher's subject area. Teachers uploaded their videos to a secure website, where they identified the topic of the lesson and had the option to upload supporting materials

Content Knowledge for Teaching: Algebra assessment. The MET project administered an assessment of pedagogical content knowledge to all teachers in the study. Those who taught ninth grade algebra took a specially-designed Content Knowledge for Teaching Algebra assessment (CKT) developed by researchers at the Educational Testing Service (ETS) and the University of Michigan (Bill & Melinda Gates Foundation, 2010b; Gitomer, Phelps, et al., 2014). The CKT assessment was administered in early 2011 and focused on teachers' pedagogical content knowledge—the specialized knowledge that enables teachers to effectively teach algebra content to students. The Algebra CKT assessment included 37 selected response items (two of which were excluded for poor

performance) and had an overall reliability of 0.77 (Gitomer, Phelps, et al., 2014). I utilized teachers' CKT scores for both sampling (see below) and to better understand the information yielded by the observational instrument.

Scores on MET project observational instruments. Trained MET project raters scored lessons on multiple established classrooms observation instruments, both subject-specific (e.g. MQI, PLATO, UTOP) and general (e.g. CLASS, FFT). All videos were scored by at least one rater, with approximately 20% of videos double-scored by a second rater for the purposes of calculating interrater agreement. Raters scored the first 30 minutes of instruction of each video-recorded lessons in segments ranging from 7.5 minutes to 15 minutes, depending on the instrument. Lesson-level scores were generated by averaging segment-level scores. Teacher-level scores were aggregated in a similar manner. I use these scores alongside scores on the algebra-specific instrument to assess the degree to which these different observational instruments provide convergent and divergent information about instructional practice.

Sample

To develop the observational instrument, I utilized the same analytic subsample as in the exploratory analysis in Chapter 2. From the full sample of ninth grade algebra teachers, I excluded teachers without CKT scores and without viewable video. This yielded a sample of 81 ninth grade algebra teachers. I ranked these 81 teachers by their scores on the CKT assessment and then binned teachers into quintiles by their CKT scores. I then chose a random stratified sample of 24 teachers, eight teachers each from the high (top quintile), mid (third quintile) and low (bottom quintile) “quality” bins.¹¹

¹¹ Sampling procedures that use the MET's extant observational instrument scores to select teachers and lessons might result in a sample of instruction that privileges the factors examined by those protocols. Assuming teaching quality is normally distributed, taking a simple random sample from eligible teachers

Past research has found that instructional quality varies sharply according to teachers' mathematical knowledge (Charalambous & Hill, 2012; Hill, et al., 2008), and by sampling in this manner I hoped to maximize the variability in algebra instruction used to develop the instrument.¹² This sampling yielded a total of 75 lessons with viewable video. This sample was used in the exploratory analysis presented in Chapter 2, which led to the development of the themes that informed the creation of the observation protocol.

Instrument Development

Like other researchers who have engaged in classroom observation instrument development (see for example, Hamre, et al., 2007; LMTP, 2011), the process for developing this instrument was an extension of grounded theory (Glaser & Strauss, 1967). Here, instrument development is informed by research literature and by the data itself, and results in the development of categories that are then tested and refined via iterative processes.

To develop codes in the instrument, I worked together with a research team comprised of 4 experienced former and current secondary mathematics teachers, two of whom are mathematics education researchers, one of whom is a teacher and instructional coach, and one of whom is a staff member on mathematics education research projects. Similar to the design process of other classroom observation instruments (see, for example, LMTP, 2011), our development of these codes was informed by the practices evident in the videos themselves, our own teaching experience, and existing literature around instructional practices in algebra. After developing the themes prevalent in the

may have resulted in an oversample of typical lessons, thus underestimating variability among teachers and obscuring important instructional practices in the tails of the distribution. Prior work with the 5-point Mathematical Quality of Instruction instrument has shown observational scores to be normally distributed.

¹² CKT scores in the sample are approximately normally distributed.

videos (for a full description of the analytic process that resulted in the major themes that guided code development see Chapter 2), we engaged in an iterative process to develop codes and a scoring guide. Each member of the research team watched a randomly assigned video in its entirety, blind to the teacher's CKT score. Observers recorded a brief lesson summary, including the topic of the lesson, a narrative description of the lesson, mathematical strengths and weaknesses of the lesson, and any salient instructional features. Each researcher also nominated key segments of their assigned video for the group to watch together for the purposes of code development. As we progressed and defined each code, researchers nominated segments that reflected the codes under development and that might help to clarify these practices across varying levels of quality. The nominated segments formed the basis of the research team's meetings and discussions. In developing the individual codes in the instrument, we repeatedly returned to the video to watch and re-watch segments of instruction to test and refine our emergent codes.

Through reviewing our narrative descriptions and continually watching video, we identified common dimensions of practice and articulated descriptors for various levels of quality, generating a scoring guide as we did so. We used these evolving criteria to score the next group of segments identified for study, while continuing to be open to new, emergent characteristics. This served the dual purpose of exploring dimensions that emerged from the data as well as refining those codes we had begun to develop. We calibrated and reconciled our scores on these later observations, bringing in relevant literature to make sense of our classifications and using the codes to test and refine the theoretical features of algebra instruction. We prioritized the development of codes that

described features prevalent in our sample, those not already captured by existing observational instruments, and those featured in the literature on algebra teaching and learning.

In scoring, each researcher independently watched and scored the segment on the developing codes, after which the team met to discuss and reconcile their scores. Through this process, the research team watched and discussed 20 segments of instruction, scoring an additional 18 segments to clarify the codes in the instrument and develop anchor score points. We reconciled scores to agreement, making adjustments to the codes as needed for clarity and to assist with future scoring. At this point, we were confident that we had outlined the codes with sufficient specificity and had seen ample variation in quality to anchor score points. We continued to watch and score video, scoring an additional 6 segments on the codes we developed. We reconciled our scores to agreement, making only small changes in the instrument at this point for clarification.

Observational Instrument

This process generated a small observational instrument called Quality of Instructional Practice in Algebra (QIPA). Like other observational instruments, QIPA is comprised of domains that articulate the core practices of interest. Each domain contains a set of dimensions, which describe specific instructional features.¹³ Lessons are divided into shorter segments and trained raters assign scores on a numeric scale using a rubric with descriptors articulating various score points (Gitomer, Bell, et al., 2014). The

¹³ In this study, I use the terms code and dimension interchangeably to describe particular, measurable features of instruction. Observational instruments traditionally articulate domains of instructional practice (e.g. Emotional Support), within which are specific dimensions or features of instruction (e.g. Positive Climate, Negative Climate, etc.) that are of sufficiently small grain size for raters to identify and assess.

algebra-specific instrument is comprised of codes at two levels: the lesson level and the individual segment level.

I now turn to a description of the instrument, focusing on the domains of instructional practice and the corresponding codes within each domain developed in this framework. There are two main domains of algebra instruction in the QIPA: Teaching Procedures and Leveraging Connections, each with four codes. The first domain, Teaching Procedures, captures the instructional features that occur when teachers engage specifically in instruction on procedures: *Making Sense of Procedures*, *Supporting Procedural Flexibility*, and *Organization in the Presentation of Procedures*. It also contains an *Overall Teaching Procedures* code. The second domain, Leveraging Connections, captures the nature and quality of the instructional features related to the connections teachers forge within and across algebraic concepts and occurs regardless of whether teachers engage in instruction on procedures. This domain features four codes: *Connecting Across Representations*, *Situating the Mathematics*, *Making Connections Between Concrete and Abstract Ideas in Algebra*, and *Overall Leveraging Connections*. In addition, the instrument includes two lesson-level codes intended to capture the format of instruction: *Teacher-Led Instruction* and *Inquiry/Exploration*. Below, I first describe the QIPA framework and scoring process, then provide examples of its use.

Domain 1: Teaching Procedures

In addition to requiring students to think conceptually about the mathematics, success in algebra likely demands student fluency in both symbolic representation and manipulation (Kieran, 2007). The degree to which mathematics instruction, and algebra in particular, should focus on concepts or procedures (or both) is a disagreement that has

raged for some time (Kieran, 2013). Yet mathematics education researchers acknowledge that the two types of knowledge are intertwined and likely develop in tandem. Conceptual knowledge can lead to the development of procedural fluency while procedural fluency can aid in the development of conceptual understanding (Rittle-Johnson, et al., 2001). Indeed, flexibility and efficiency with procedures may be an important aspect of algebra in particular, as the subject contains a significant amount of procedures involving symbolic manipulation (Kieran, 2013). Furthermore, procedures are not only prevalent in algebra, but they are often more complex than those in arithmetic (Star, 2005; Star & Rittle-Johnson, 2009). In fact, deep procedural knowledge—using procedures flexibly and efficiently—may be highly beneficial to student understanding in algebra (Star, 2005). When thinking about algebra, it may in fact be more useful to think about its symbolic procedural activities *in terms of concepts*. In algebra, procedures may be conceptual in nature—both in their initial introduction and as they are revised to encompass new mathematical ideas and situations (Kieran, 2013). Little research has illuminated instructional strategies that specifically improve conceptual understanding (Hiebert & Grouws, 2007) and it may be that high quality instruction on procedures with an attention to the concepts underpinning them does just that.

Thus, I developed a set of codes designed to identify particular features of instruction that teachers engage in *when teaching procedures*. While instruction varied in quality along these particular aspects, it may be that each element, alone or in combination, benefits student learning. I define procedures here as the instructions for completing a mathematical algorithm or process. Instruction on procedures encompasses the presentation of new procedures, the review of previously learned procedures and the

descriptions of procedures in the context of solving problems. These codes are intended to capture the teaching of procedures in multiple contexts—such as during direct instruction, while a teacher is interacting with students during independent or group work time, or when the teacher (or students) are presenting solutions to previously worked examples.

The QIPA contains three dimensions to capture the quality of the instruction on procedures in a lesson (for a full list including detailed descriptions of each code see Table 2). The first code (*Making Sense of Procedures*) captures the degree and depth with which teachers and students make sense of procedures, either by attending to meaning of the individual steps of the procedure, the solution generated by the procedure, or to the procedure as a whole. The next code (*Supporting Procedural Flexibility*) focuses on the degree to which teachers present algebraic procedures to students in ways that afford students the opportunity to develop procedural flexibility. For example, this code captures such practices as noting multiple pathways through a procedure, noting when a particular procedure is appropriate or what cues the selection of that particular procedure, and comparing multiple procedures for their affordances or limitations. The third code in this domain (*Organization in the Presentation of Procedures*) indicates how complete, detailed, correct, and organized the teacher (or students') presentation of content is when describing or outlining a procedure. Even in the absence of making meaning around procedures or other nuances, at a minimum we were interested to see whether and to what degree when teachers present a procedure, the steps of that procedure are correct, explained clearly—either verbally or in writing—and are organized in a way that afford students the opportunity to replicate the procedure on their own.

Domain 2: Leveraging Connections

While teaching procedures was a frequent occurrence in the sample, there were also instructional features that emerged both in the context of teaching procedures and outside of it. Specifically, I noted multiple ways in which teachers aimed to make connections between and within mathematical ideas. These connections may be important because they provide the opportunity to give mathematical meaning to algebraic abstraction and thus have the potential to attend to student difficulties with algebra. The abstract nature of algebra and the prevalence of symbolic representations create difficulties for students. In addition, there are also inflection points *within* the algebra curriculum that require students to adapt and expand their prior understandings. Students may face difficulties transitioning between topics in algebra, particularly when topics are related but rise in complexity, requiring students to adapt their understanding of particular processes to new situations (Kieran, 2013).

One way in which teachers attend to these issues is by making connections for students that aim to bridge the mathematical divide between seemingly disparate topics, between abstract algebraic ideas and their concrete underpinnings, and across representations that convey mathematical relationships. Thus, the instrument contains three dimensions intended to capture the nature of these connections that teachers leverage to build student understanding in algebra (for a full description of each code see Table 3). The first code in this domain (*Connecting Across Representations*) captures the nature of the connections teachers and students make between and across representational forms in algebra—specifically between graphs, tables, equations, and problem contexts. Connections across representations are an important way to develop meaning in algebra

(Knuth, 2000) and this practice has been encouraged in recent decades of mathematics reform. The second code (*Situating the Mathematics*) looks at the connections teachers and students make across aspects of the algebra curriculum, to related topics, or to the broader domain of mathematics. Teachers do this in a variety of ways, including making connections between a current topic to prior or future content in algebra, using the architecture of the lesson to develop the mathematics through the course of a single lesson, or to connect what students are learning to the broader domain of mathematics. In doing so, we speculate that teachers lay groundwork to motivate the current topic under study within a broader mathematical context. Finally, the third codes in this domain (*Making Connections Between Concrete and Abstract Ideas in Algebra*) is intended to capture the degree to which teachers and students leverage concrete examples, representations, or ideas to develop understanding of abstract concepts, formulas, notation or definitions. In making linkages between abstract or generalized ideas and their concrete underpinnings, teachers may build bridges for students in developing understanding of abstract ideas. Teachers do this, for example, by attending to what the components of a mathematical formula represent, clarifying mathematical definitions or abstract algebraic concepts with concrete examples (or non-examples), or explicitly relating an abstract concept to its analogous concrete idea.

Whole-Lesson Instructional Format

In addition to specific instructional practices, which were measured at the segment level, I was also interested in the degree to which whole lessons followed specific instructional formats. Specifically, I was interested in the degree to which lessons were teacher-led and featured students actively engaging in the mathematics, as well as

the degree to which lessons contained mathematical inquiry or explanation. To address this, I developed two whole-lesson codes for the QIPA designed to capture the broader format of the lessons. The first code (*Inquiry/Exploration*) was designed to capture the degree to which students were asked to do significant mathematical work involving mathematical investigation or discovery. For example, lessons that engaged in inquiry or exploration included an extended, open-ended investigation without a pre-determined solution path in which students work together to explore the mathematics with limited teacher direction. In addition, we developed a second code (*Teacher-Led Instruction*) intended to capture the degree to which the teacher directed the mathematical content and processes in the lesson and the degree to which students were actively engaged in doing mathematics in the lesson. These codes are not necessarily in opposition to one another, as they focus on duration of these instructional formats. Thus, it is possible that a lesson contain some time in which students engage in inquiry and exploration and some time in which the teacher is heavily directing the content. In addition, there are other lesson formats (e.g. completing test review worksheets in stations) that may not fit into either of these two categories. As such, these codes are not intended to be exhaustive of all possible lesson-formats, but rather to capture the degree to which particular formats noted in the literature are present in algebra lessons.

Scoring Algebra Lessons

As we finalized the dimensions of the instrument, we also developed a rubric for scoring instruction across levels of quality. Consistent with other observational instruments (e.g. Gitomer, Bell, et al., 2014; Hamre, et al., 2007; Hill, et al., 2008; LMT, 2011), we scored lessons by breaking instruction into segments. Like Hill and colleagues

(2014), we broke instruction into 7.5-minute segments, as this was a short enough period of time to keep track of mathematical practices and events, but sufficiently long to view instructional interactions happen in the broader context of the lesson and to score efficiently. Raters watched a 7.5-minute segment and then paused the video to record their scores. After each segment, raters indicated whether procedures were taught in that segment. If procedures were taught, the raters scored the segment on the Teaching Procedures codes. If procedures were not taught, these segments were automatically scored at the lowest score point.¹⁴ Raters scored all segments on the Leveraging Connections codes. At the completion of the lesson, raters scored the entire lesson on the lesson-level format of instruction codes. Raters scored instruction on a scale of Low (1) to High (5). A score of Low (1) indicates that the practice captured by the code did not occur or the teacher engaged in the practice but it was mathematically incorrect. A score of Mid (3) was anchored to indicate “modal practice.” For each code, this indicates more than brief attention to the particular practice in the segment but without depth or elaboration. A score of High (5) indicates that the particular instructional feature characterized the segment and was instantiated at high levels of quality. Scores of two and four were used for cases that fell between Low and Mid or Mid and High, respectively. Segments were scored two if the particular instructional feature was briefly present—even in passing. Segments were scored four if the instructional feature was more developed than a Mid score would indicate, but did not reach the threshold for a

¹⁴ Because a score of Low indicates an absence of a particular feature, I wanted to distinguish segments in which no procedures were taught (and thus scored Low on the codes in the Teaching Procedures domain by definition) and those in which procedures were taught but the teacher did not engage in the practices outlined in the codes. As a result, each segment was also scored for whether instruction on procedures was present or not. For the purposes of analysis, this allows me to discuss the quality of instruction on procedures *when procedures are taught*.

High score. For a complete scoring guide, see Appendix B. Whole-lesson codes (*Inquiry/Exploration* and *Teacher-Led Instruction*) were scored in a similar manner.

After completing each lesson, raters gave the whole lesson a score ranging from Low (1) to High (5), indicating the degree to which instruction was teacher led or the lesson featured inquiry or exploration.

Examples of Scored Instruction

To illustrate how the scoring rules were instantiated, I present three examples of instruction and discuss the scores on each. In the first example, the segment scored Low on most codes in the QIPA. The second example illustrates instruction that scored a Mid on *Making Sense of Procedures*. Finally, the third example demonstrates instruction that scored High on *Supporting Procedural Flexibility*. For each example, I give a brief, general description of the segment and note the ways in which the teacher engaged (or did not engage) in the instructional features outlined in particular codes.

In a lesson quadratic functions, the teacher explained to students that the goal of the lesson was to convert quadratic functions from vertex form [$y = (a(x - h)^2 + k)$] to standard form [$y = (ax^2 + bx + c)$]. In one particular segment, she asked students to think about and discuss how the two forms of quadratic equations were different. While the question offered students the opportunity to discuss the roles of the various coefficients in the two equations and what information each reveals about quadratic functions, the ensuing discussion focused on superficial details of how the two forms differed structurally. One student noted, for example, that the a is on the outside of the parentheses in vertex form and attached to the x^2 in standard form. Another student observed that one equation has h 's and k 's and the other has b and c . The teacher

acknowledged these observations and told students that they needed to be able to manipulate an equation in vertex form so that “it looked like an equation in standard form.” She then demonstrated the procedure for doing so, but did not discuss with student the purpose of the procedure she was teaching. The remainder of the segment focused on the symbolic manipulation she demonstrated absent any meaning, context, or connection to the graphical representation of quadratic functions. Thus, this segment scored Low on a number of codes, including *Making Sense of Procedures*, *Connecting Across Representations*, and *Situating the Mathematics*.

In many lessons, teachers engaged in instruction that featured the instructional practices measured by the instrument to some degree—for a small part of the segment or without much depth. For example, a segment that scored Mid on *Making Sense of Procedures* might include a short explanation of why a step in a procedure holds mathematically, but not include elaboration or much detail. In one segment typical of those that scored Mid on this code, a teacher presented the example $\frac{x}{\sqrt{2}}$ and explained to students that it needed to be simplified, stating:

There’s a rule that says you can’t have a radical in the denominator... you can’t have a radical, a square root on the bottom. It’s okay to have a square root at the top... I’m not allowed to have a square root sign at the bottom so I have to fix it. I have to get that two out of there and here's how I do it. You multiply... by square root of two over square root of two. You take whatever number is in the denominator and you write it twice.

This part of the procedure was presented simply as a series of steps with no attention to meaning. However, the teacher then continued:

So let’s talk about why. First of all, what does the square root of two over the square root of two equal?... Anytime you have a number and the identical number underneath it, what does this number right here equal?... One. And so I’m multiplying by one. If I’m multiplying by one does it

change anything? ... If I multiply by one, it doesn't really change my equation, it only changes the way my expression looks.

In this part of her explanation, the teacher reinforced the mathematical reason behind this particular step in the procedure, reinforcing the concept that multiplying an expression by one does not change its value. The teacher continued in this vein, engaging in a one-on-one explanation with a student who was confused about why she had chosen to multiply the expression by $\frac{\sqrt{2}}{\sqrt{2}}$. While her explanation was not sustained nor in depth, it was developed more than briefly and thus this segment received a score of Mid for *Making Sense of Procedures*.

In contrast, a score of High (5) indicates that the particular instructional feature characterized the segment either by a sustained focus on that particular feature or through an extended and well-developed instance of that particular feature. For example, in a lesson reviewing the procedure for solving systems of equations, a teacher placed strong emphasis on supporting flexibility, with one segment scoring High on this code. During this segment, the teacher discussed and compared the multiple ways of solving systems of equations and emphasized what aspects of particular problems cued the selection of one particular method over another. The teacher repeatedly told students that the goal of reviewing the methods was to determine which method makes sense given the structure of the system presented. After reviewing the procedures for two different methods for solving systems of equations (substitution and elimination), the teacher directed students to look at the structure of the equations themselves and think about what method would work best given how the equations are written. The teacher acknowledged that it was possible to manipulate the equations so that any method could be used, but that the process of doing so might complicate the problem or lead to error. She told the students:

What did I say the theme was today? Best method. Best method. Say it again, best method. If they give me [an equation] in y equals [form], am I going to rewrite it in standard form? If they give it to me in standard form, am I going to rewrite it in y equals form? That's what you did on the assessment. You got good at one of these (methods) and tried to use it for everything. But that didn't go so well.

She went on to emphasize to students that the way in which the problem is written helps to cue the best, or most efficient, method for solving the problem. To further underscore this, the teacher then gave three different systems of equations and rather than asking students to solve each, she asks which method students would choose and why. This emphasis meant that for the entire segment, a sustained focus of the instruction was on building students' procedural flexibility and thus this segment of instruction scored High on this particular code.

Validity

If the purpose of an observational protocol is to improve instruction, then it is important to investigate the information provided by the scores on the instrument (Gitomer, Bell, et al., 2014). While the validity and reliability concerns for an instrument designed to evaluate teachers are different than those for one intended to describe instruction and provide a common language for instructional improvement, it remains important to examine the quality of the information yielded by the instrument. Investigations of validity often include examining scoring inferences (such as the degree to which the scores accurately reflect the interactions in the classroom and the consistency with which the scoring rules are applied), the generalizability of inferences from the particular sample observed, the degree to which scores are related to other measures of teaching quality, and the appropriateness of the connection between the observation scores and the interpretations/use of those scores (Bell, et al., 2012). While a

full validity argument (see Kane, 2006) is outside the scope of this study due to sampling design, I investigated the quality of the information produced by the codes developed in the algebra-specific instrument.¹⁵ I first scored all videos in the subsample described above (n=73 lessons) on the QIPA blind to the teacher's score on the CKT assessment.¹⁶ The order of teachers and the order of lessons were randomly generated. In what follows, I report on the reliability of the instrument and the validity of the inferences that can be drawn from scores on these codes.

Consistency and Accuracy

Following Bell and colleagues (2012), I assessed the degree to which the scoring rules were applied sufficiently accurately and consistently. I first examined the descriptive statistics for each code, assessing the degree to which observations were scored across the full range of score points; given that score points were developed from this sample, all should be used in the analysis. In Table 4, I show the mean, standard deviation, minimum and maximum scores across all segments for each code. For the codes in the Teaching Procedures Domain (*Making Sense of Procedures*, *Supporting Procedural Flexibility*, and *Organization in the Presentation of the Procedures*) this includes all segments in which procedures were taught (n=261). For the remaining codes, this includes all segments from all lessons (n=465). I find that at the segment level, observers utilized all score points, although scores were not evenly distributed across all

¹⁵ Ideally, a G-Study would be appropriate for this purpose. However, I do not have a study design that is adequate for this purpose. For example, a full g-study necessitates a large sample with multiple raters that is beyond the scope of this study.

¹⁶ In the scoring process, two lessons were excluded from analysis. In one lesson, the teacher conducted instruction entirely in Spanish and non-Spanish-speaking raters did not feel that they would be able to accurately or reliably score instructional practice on the instrument. In another lesson, students worked independently on computers for the duration of the lesson and the teacher did not interact with students. As it was impossible to determine the content or activities in the lesson, and thus impossible to score instructional practice, this lesson was also dropped from the analysis.

score points. This is not necessarily a cause for concern, as it is reasonable to expect that teachers may engage in these practices in some segments more than others. At the lesson level, I find that observers similarly utilized all score points.

In Figure 4, I show the distribution of scores on each code across all score points at the segment level. Here too it is clear that scores are not evenly distributed. In particular, few segments score at High levels on any of the codes. At the segment level, scores for the *Making Sense of Procedures*, *Connecting Across Representations*, *Situating the Mathematics*, and *Connecting Between Concrete and Abstract Ideas in Algebra* dimensions are clustered at the Low score point. Previous studies of math classrooms find particular practices related to cognitive challenge and sense-making instantiated rarely or poorly (e.g. Gitomer, Bell, et al., 2104; Kane & Staiger, 2012). This may be the case with this instrument as well. It may also be reasonable to expect that these instructional features not appear in every segment of instruction, and as such have skewed distributions at the segment level.

Aggregating scores to the lesson level, I find the distribution of scores looks somewhat more normally distributed for most codes, with the notable exception of *Situating the Mathematics* (See Figure 5).¹⁷ This practice appears to be clustered in a small number of lessons when it appears. This may reflect the relative absence of this particular practice or it may reflect scoring error. Scores on this code in analysis from other samples also shows this practice to be relatively rare, supporting the idea that this clustering may not be due to scoring error. In addition, lesson-level agreement rates for

¹⁷ Like Hill and colleagues (2014), I calculated average scores for each code at the lesson-level, shrunk for the number of segments per lesson to address the fact that some lessons contained more segments than others. I calculated teacher-level scores for each code in a similar manner, using shrunk averages to account for the fact that some teachers had more lessons than others.

this code were within one score point on double-scored lessons over 90% of the time (see discussion below), providing more evidence that this clustering is likely not due to scoring error. The distributions of the remaining codes in the Teaching Procedures domain indicate that there is variation in quality at the lesson-level, with most lessons clustering around the mean and fewer lessons at the extremes. The lesson-level codes in the Leveraging Connections domain are not as normally distributed. Again, this may not weaken the inference if the purpose of the instrument is to be used with teachers as a tool for instructional improvement. If we believe these instructional features to be important and beneficial for students, then it may be less consequential that scores cluster in particular score-points and may reflect actual practice rather than instrument design issues. Again, interrater reliability evidence reflects that raters were within one score point on these codes for between 90% and 100% of the lessons, providing convergent evidence that this clustering may not be due to scoring error.

Another way to address consistency in the instrument is to investigate the degree to which particular practices vary across lessons taught by the same teacher versus across teachers. In Table 5, I present the percent of variation in a given teacher/lesson/segment score combination that is attributable to variation within teacher, across lessons (Column 1) and across teachers (Column 2), calculated using intra-class correlations. These results indicate that there is greater variability across lessons within teachers than across teachers for *Organization in the Presentation of Procedures*, *Connecting Across Representations*, *Situating the Mathematics*, *Connecting Between Concrete and Abstract Ideas in Algebra*, and *Teacher-Led Instruction*. Thus engaging in these particular instructional features and format appear to be more lesson- than teacher-dependent. In the case of *Connecting*

Across Representations, this is not surprising as the opportunity to engage in this instructional feature may be more dependent on the content of a particular lesson. For example, a lesson on linear inequalities may offer teachers more opportunities to connect across representations than a lesson on the rules of exponents. In the case of *Organization in the Presentation of the Procedures*, this difference may reflect teacher decision-making around when to be more explicit and detailed (e.g. introducing material for the first time rather than reviewing previously taught content), which would also be lesson-dependent. In contrast, there is more variation between teachers than within teachers across lessons for *Making Sense of Procedures*, indicating that this practice may be less dependent on lesson context or content. Understanding the sources of variation in the scores is useful when thinking about the inferences drawn from particular scores.

Reliability of Scores

For any protocol aiming to describe instruction or provide feedback to teachers, it is important to know the degree to which independent observers watching the same instruction would rate it similarly using the protocol. Lack of agreement creates a challenge toward instructional improvement because it may be a signal that observers see practices differently. To investigate the reliability of scores on the instrument, I randomly selected 20% of videos to be double scored by trained raters for the purposes of interrater reliability, resulting in 15 double-scored lessons. I investigated observers' agreement with one another, calculating the degree to which the scoring rules were applied consistently across raters. In Table 6, I report the percent agreement (the percent of segments in which gave the same score), percent adjacent agreement (the percent of segments in which raters were off by at most one score point), and percent disagreement (the percent of

segments in which raters were off by greater than one score point). Agreement rates are relatively high compared to other classroom observation instruments (see for example Gitomer, Bell, et al., 2014) and adjacent agreement rates suggest that raters were within one point of each other over 90% of the time for the dimensions in the Teaching Procedures domain and over 85% of the time for the dimensions in the Leveraging Connections domain. Weighted kappas suggest that the inter-rater reliability ranges between 0.37 and 0.65, quite high compared to other observation instruments used with in the MET project study (Kane & Staiger, 2012). Reliabilities were lower at the segment level for the *Making Sense of Procedures* code relative to the other codes in the Teaching Procedures domain. Similarly, reliabilities for the codes in the Leveraging Connections domain were lower relative to those in the Teaching Procedures domain. Other research finds that observers are more likely to agree on certain elements of instruction—particularly those that can be directly observed—than on instructional features that require more inference (Gitomer, Bell, et al., 2014). It may, for example, be easier for raters to identify the qualities of organization in teaching procedures than to identify sense-making in the context of teaching procedures.

In Table 7, I present similar information at the lesson level. To calculate agreement rates at the lesson level, I averaged scores across segments and rounded these average scores to the nearest score point to generate scores for comparison across raters. I present only exact agreement and disagreement rates here, as I find that raters were within one score point from each other on all lessons on all codes except for *Situating the Mathematics*. Weighted kappas show high levels of agreement on some codes, while for other codes (*Making Sense of Procedures*, *Supporting Procedural Flexibility*, and

Situating the Mathematics), reliabilities decrease from the segment level. This may be due to small sample sizes and scores on these particular codes being constrained to only a few score points. For example, by design Cohen's kappa is lower when distributions are more skewed. For example, the distribution of the *Situating the Mathematics* code at the lesson level (see Figure 5) indicates that segments that are scoring above low on this code are clustered in few lessons. However, it is worth noting that lesson-level reliabilities are moderate to high compared to other observation instruments with two raters (Kane & Staiger, 2012).

Taken together, these results indicate that scores on the QIPA from two raters are frequently within one score point and that the individual codes in the instrument have moderate reliability. While this level of reliability would be questionable if the protocol were to be used for evaluation purposes, it may be sufficient to allow for instructional feedback. For example, teachers and observers working together to understand and use observation protocols may improve both reliability as well as instructional practice through the process of reconciling scores to agreement. However, because scores in some dimensions were clustered at the low end of the rating scale, it may be necessary to investigate inter-rater reliability in larger samples.

Convergent and Discriminant Validity

Another important aspect of an investigation of the validity of an observational instrument is to assess convergent validity, or the degree to which the information provided by the instrument aligns with other measures of teacher and teaching quality. For example, the codes I have developed are potentially related to (and a product of) teacher knowledge and are a part of the complex set of classroom interactions that

contribute to student achievement (Cohen, et al., 2003). Thus it is important to investigate the degree to which the codes I have developed correlate with other measures of teaching quality such as teacher knowledge and those measured by other observational instruments. In Table 8, I present estimated Pearson correlations between teacher-level scores on the dimensions and teachers' scores on the CKT assessment administered by the MET project.¹⁸ I find that at the teacher level, CKT score is positively correlated with the whole-lesson code of *Inquiry/Exploration* (0.6113, $p < 0.001$), indicating that teachers with higher CKT scores were associated with engaging in more investigation and exploration. The *Teacher-Led Instruction* code is negatively correlated with CKT score (-0.4465, $p < 0.05$), indicating that teachers with higher CKT scores scored lower on average on this particular code, meaning that instruction tended to be less teacher-driven and more student-centered. This is consistent with other studies that find correlations between teacher knowledge and more reform-oriented instructional formats (e.g. Hill, Rowan, & Ball, 2005; LMT, 2011).

Looking at the relationship between instructional features and teachers' CKT, the results are mixed. Teachers' CKT scores are positively correlated with *Connecting Between Concrete and Abstract Ideas in Algebra* (0.5182, $p < 0.05$). This indicates that teachers with higher CKT scores scored higher on average on this particular code. Surprisingly, there is a negative correlation between teachers' CKT score and *Organization in the Presentation of the Procedures* (-0.4351, $p < 0.05$). This would indicate that instruction from teachers with higher CKT scores rated lower on this particular instructional feature. This correlation could indicate that this particular practice

¹⁸ One teacher in the sample, who only had one video-recorded lesson available, taught entirely in Spanish. This lesson (and thus this particular teacher) was excluded from the teacher-level analyses, yielding a sample size of 23 teachers for teacher-level analyses.

may be negatively related to teacher pedagogical content knowledge or may indicate that the CKT assessment is not designed to measure this type of instructional feature and the statistically significant results may be a function of small sample size and measurement error. Indeed, an examination of the scatter plot of CKT and scores on this code show that there is an atypical teacher with an unusually low CKT score and a relatively high score on this particular code. With such a small sample size, this one teacher may be exerting disproportionate influence on the correlation. Without this teacher included, the magnitude of the correlation is greatly reduced and it is no longer statistically significant. Teachers' CKT is not correlated with the other instructional features measured by the instrument. However, these results should be interpreted with caution given the small sample size ($n = 23$ teachers). Larger samples of teachers may unearth relationships that I am unable to detect here.

The results presented in Table 8 also display the correlations between dimensions of the QIPA at the teacher level. Not surprisingly, the *Overall Procedures* code is positively correlated with the three codes in the Teaching Procedures domain. This is to be expected as it is intended as a holistic assessment of these component features. Similarly, the *Overall Connections* code is positively correlated with two of the three codes in the Leveraging Connections domain (while there is no statistically significant correlation between *Overall Connections* and *Situating the Mathematics*, this may be a function of both sample size and the relative absence of this particular instructional feature). Looking at the relationship between individual codes, I find that *Organization in the Presentation of Procedures* is highly correlated with *Supporting Procedural Flexibility* ($0.7638, p < 0.001$), indicating a relationship between these two features at the

teacher level. *Connecting between Concrete and Abstract Ideas in Algebra* is also positively correlated with *Making Sense of Procedures* (0.4734, $p < 0.05$) and *Connecting Between Representations* (0.6992, $p < 0.001$). This may indicate that teachers who engage in this practice also engage in other sense-making practices in their instruction.

Finally, I drew on the MET classroom observation scores to investigate the degree to which scores on the dimensions measured by the QIPA are correlated with scores on codes from relevant existing instruments in order to assess whether the algebra-specific codes may be capturing potentially distinct aspects of instructional practice from those seen in extant protocols. In Table 9, I present Pearson correlations between the codes on the QIPA at the lesson-level and lesson-level scores on the MQI-Lite from the MET study for those lessons from this sample with MQI scores ($n=53$). I find no statistically significant correlations between some dimensions of the QIPA and the dimensions of the MQI. Specifically, the codes in the Teaching Procedures domain (*Making Sense of Procedures*, *Supporting Procedural Flexibility*, and *Organization in the Presentation of Procedures*) are uncorrelated with all of the MQI dimensions, indicating that the algebra instrument may be measuring something distinct in the domain of Teaching Procedures.

There are, however, some statistically significant correlations between scores on the Leveraging Connections codes and some of the dimensions of the MQI. Specifically, *Connecting Across Representations* is correlated with the Richness dimension of the MQI (0.4006, $p < 0.01$). This is not surprising as this particular instructional feature (defined somewhat more broadly) is a component of the Richness dimension of the MQI. *Situating the Mathematics* is also correlated with the Richness dimension (0.4135, $p < 0.01$), as well as with the Student Participation in Meaning Making and Reasoning dimension of

the MQI (0.3576, $p < 0.01$). This may be a function of the fact that this particular code is closely aligned with the reform-oriented practices the MQI is aiming to capture.

Discussion

In this paper, I described the development of an observational instrument geared to the features and format of instruction in algebra specifically. The goal of creating the QIPA was to provide a way to systematically describe the nature and quality of instruction in a large sample of algebra lessons. Doing so also provides a potential tool for researchers and others engaged in the work of instructional improvement. Standardized measures of instructional practice provide both a common language and framework for discussions about instructional practice (Gitomer, Bell, et al., 2014; Pianta & Hamre, 2009) and an opportunity to further investigate the effectiveness of various interventions on instructional quality (LMT, 2011). Other research might consider the relationship between these particular practices and student learning outcomes.

Prior analysis (see Chapter 2) yielded themes around instructional formats and features that were then codified into an observational protocol. Specifically, the protocol aims to capture instructional features related to the domains of teaching procedures and leveraging connections in the context of the algebra classroom. The dimensions captured by the QIPA in the Teaching Procedures domain include *Making Sense of Procedures*, *Supporting Procedural Flexibility*, and *Organization in the Presentation of Procedures*. The dimensions captured in the Leveraging Connections domain include *Connecting Across Representations*, *Situating the Mathematics*, and *Connecting Between Concrete and Abstract Ideas in Algebra*. In addition, the instrument captures instructional formats in algebra lessons, measuring the degree to which the lessons are teacher-directed and

the degree to which students engage in mathematical inquiry or exploration. These are certainly not the only instructional formats and features that might be measured by an algebra-specific instrument. Rather, these dimensions should be considered the beginning of an engagement in the instructional features inherent in algebra classrooms and reflective of practices that emerged from an analysis of contemporary instruction.

Investigating the measurement properties of the QIPA, I find evidence that it measures distinct aspects of instructional practice from other mathematics-oriented instruments, although there is some crossover. For example, the algebra instrument converges with the MQI on more reform-oriented instructional features such as connecting between mathematical representations. However, it appears that the codes in the Teaching Procedures domain are a unique contribution of this instrument as they do not appear correlated with the codes on the MQI. The reliability of the individual codes of the QIPA are high relative to other observational instruments (see for example the discussion in Kane & Staiger, 2012). However, similar to other validity studies, I find that those instructional practices that scored lowest on average are some of the most difficult for raters to agree on (Gitomer, Bell, et al., 2014; Kane & Staiger, 2012).

Taken together, these results indicate that attention must be paid to the uses of an instrument such as QIPA and the inferences drawn from scores. The reliability of the scores on particular codes may not support using the instrument for a summative evaluation of teachers' practices, and certainly not in a high-stakes manner. In addition, the domains measured by the instrument are certainly not the only instructional practices in which we would hope algebra teachers engage. As I am not aiming in this study to measure teacher effectiveness (or teaching effectiveness per se), but rather to conduct a

descriptive analysis of the nature and quality of algebra instruction, the instrument remains useful for this particular purpose. It does appear that the QIPA is capturing instructional features that are distinct from those measured by other mathematics-oriented instruments, particularly in the area of teaching procedures. As such, it can serve to open up discussion and understanding of both the component aspects of teaching procedures and what it might look like to engage in high quality instruction on procedures. Furthermore, an observation protocol such as QIPA that focuses on more fine-grained instructional features when used as a tool for teacher learning and instructional improvement efforts may provide common language for teachers around how to engage in these practices at high levels of quality.

Future research is needed to apply these codes to a larger sample of algebra lessons. Though teachers' scores on the CKT assessment are correlated to some degree with particular codes in the algebra instrument, small sample size may mask some associations. Additional validation work would also correlate scores on the algebra instrument with student outcome measures so as to assess the degree to which the particular practices measured by the instrument are related to, and may positively impact, student learning in algebra.

Given the evidence and cautions discussed above, it is important to discuss potential uses of such an instrument. Traditionally, measures of the quality of instruction focus indirectly on student achievement and other measures of student learning. Yet these methods are frequently critiqued by practitioners and administrators for their lack of guidance on how to improve teaching (and learning). Indeed, the information generated by such measures as value-added scores or student growth percentages are not designed

to provide specific suggestions for instructional improvement. Increasingly, classroom observation instruments are included in measures of teacher and teaching quality for related evaluative purposes, with the hope that these measures *do* begin to provide feedback. Yet the instruments used most frequently in teacher accountability systems are necessarily focused on broader domains of instruction such as emotional climate, differentiation, or classroom management. Some of these domains have proven more reliable to measure than those associated with instructional practice (Gitomer, Bell, et al., 2014), an important criteria if the use of the rubric is for teacher evaluation with potentially high-stakes consequences. Teaching is a multi-layered and complex endeavor. Given the specificity of the features of instruction, it is likely unfair (and largely irresponsible) to evaluate teacher effectiveness using such an instrument.

An observational instrument focused on more fine-grained instructional features thus may not be useful for evaluative purposes in a teacher accountability framework. But by its explicit focus on discipline-specific instructional practices, such an instrument may provide more leverage as a tool for instructional improvement. For example, such an instrument has utility as a framework for grounding discussions around instruction and subsequent planning and teaching. It can provide for teachers both a common language and structure for teachers engaged in instructional improvement efforts (Gitomer, Bell, et al., 2014). By watching and scoring video clips of instruction (even instruction not their own) on an instrument such as QIPA and discussing their scores, algebra teachers have the opportunity to engage in discussions of improving practice grounded in the artifacts of teaching (Gitomer, Bell, et al., 2014). This allows teachers to begin to build a shared understanding of these practices, enabling explicit planning for incorporating particular

instructional features into their own instruction. In addition, professional development designed with the use of such an instrument can create experiences that are grounded in algebra content, engage teachers in active learning, and are coherently structured around a common framework, all hallmarks of what we understand to be effective professional development (Desimone, 2011). The instrument itself also provides a mechanism by which to evaluate the effectiveness of the professional development intervention.

The QIPA thus provides an opportunity to develop a better understanding of existing instructional practices in algebra by going beyond instructional format to investigate the nature and quality of particular instructional features. Given that these practices appear to be ones in which teachers *already* engage, teachers may be able to recognize these features in their own instruction and work to deepen them. Scores generated by the protocol have the potential to open a discussion about the nature and quality of particular instructional moments. In working with teachers to deepen their discipline-specific practices, such an instrument provides more detailed and fine-grained information regarding instructional practice. Furthermore, by providing teachers such an instrument as a tool for their own development, we can support efforts to ground instructional improvement firmly in the work of teaching.

Chapter 4.

The Nature and Quality of Algebra Instruction in Five Urban Districts

Introduction

In recent decades, policymakers have called for increased proficiency in mathematics among American students (Brown, et al., 2013; Gardner, 1983; President's Council of Advisors on Science and Technology, 2010; Tate, 1997). As a result, federal, state, and local policies have attempted to increase the amount and rigor of mathematics in schools. In particular, access to algebra has been a focus of many of these efforts (Allensworth, Nomi, Montgomery, & Lee, 2009; Clotfelter, Ladd, & Vigdor, 2012; Cortes, Goodman, & Nomi, 2015; Gamoran & Hannigan, 2000), as a body of descriptive research suggests that algebra is a key determinant of future academic success (Adelman, 2006; Chazan, 2008; Stein, et al., 2011). Algebra serves as a gatekeeper course for higher-level mathematics and, as such, has solidified its place in the research and policy spotlight.

During this time, as more students have enrolled in algebra courses in U.S. schools, a vision of *what* should constitute the domain of school algebra has begun to coalesce (NCTM, 2000; NGA, 2010). With this has also come a push for more ambitious mathematics teaching (NGA, 2010)—in algebra specifically and in mathematics more generally. Researchers and policymakers alike advocate for classroom lessons to be structured to include opportunities for students to make meaning of mathematical concepts and engage in high cognitive demand tasks, as well as to emphasize particular mathematical practices such as explanations, reasoning and justification, and mathematical modeling. Yet although these structures and practices have been

increasingly incorporated into recent standards and curriculum, there is little descriptive research on how, or even whether, this vision for algebra manifests in classrooms. While there are case studies of algebra classrooms and the quality of the instructional practices therein (e.g., Boaler, 2002; Brown, 2004; Chazan, 2000), large-scale examinations of instructional practices in algebra are rare. In fact, the *qualities* of current instruction are largely absent from the broader policy conversation.

In this study, I investigate ninth grade algebra instruction across a large sample of video-recorded classroom lessons. Drawing on data from the Measures of Effective Teaching (MET) Project and using an observational instrument specific to algebra, I ask the following: *To what degree and at what level of quality do algebra-specific features of instructional practice occur in lessons across five urban districts? What is the nature and quality of instruction that features these practices?*

Below, I ground this study in a discussion of what is known about algebra instruction and the measurement of instructional quality. I next discuss the data and sample used in this study and introduce the algebra-oriented observational instrument used to measure instructional practice. After describing the analytic strategies used in the subsequent analysis, I present both descriptive findings from scored lessons and illustrative cases designed to more clearly articulate the nature and quality of algebra instruction in this sample. Finally, I discuss the implications of these findings.

Background/Literature Review

Instructional Practice in Algebra

In the United States, algebra was traditionally conceptualized of and taught as “generalized arithmetic,” with a focus on symbol manipulation in expressions and

equations (Kaput, 1999; Kieran, 2007). Over time, scholars worried that this traditional focus on procedures around symbolic manipulation (particularly in the absence of a focus on the concepts underlying those procedures) would inhibit students' deep understanding of mathematics (e.g., Carpenter & Lehrer, 1999). Indeed, scholars argue that while a solely procedural orientation allows students to become proficient in *some* areas of algebra, when these skills are developed in the absence of understanding the mathematics behind the procedures or reasoning about solutions, students' understanding is "fragile" (Hiebert, 1999, p. 12). This tenuous understanding creates impediments to such higher-order cognitive processes such as transfer of learning to new contexts (Skemp, 2006). Such an approach also encourages students to focus on memorized procedures and to view algebra as "operations on strings of symbols" (Kaput, 1999, p. 133) rather than as a study of mathematical relationships.

Arguing that procedural fluency is not sufficient, researchers in recent decades have instead pushed for algebra instruction to support a meaning-making orientation—one that includes making meaning of procedures—and have emphasized the adoption of classroom practices designed to support and promote conceptual understanding (Kilpatrick & Izak, 2008). Modern standards have evolved to encompass this view, increasingly emphasizing meaning around concepts over simple skill mastery (NCTM, 1989, 2000; NGA, 2010). This has resulted in a de-emphasis on (though not an erasure of) paper and pencil computations and a push for instruction focused on mathematical meaning. For example, standards documents have increasingly advocated that mathematics teachers engage in such practices as using multiple strategies to solve problems, allowing opportunities for students to provide explanations and justify their

mathematical reasoning, increasing the amount and depth of student discourse around mathematics in classrooms, making connections between mathematical topics, increasing the use of real-world, contextualized problems, and using multiple representations (such as tables, graphs, and equations) to connect mathematical ideas (Kieran, 2007; NCTM, 1989, 2000; NGA, 2010).

These standards have also emphasized student participation in and ownership of mathematical learning, suggesting that students take on a more active role in the learning process and be encouraged to communicate their mathematical thinking (McCaffrey, et al., 2001). This stance is amplified by the new Common Core requirements. In particular, the new standards for mathematical practices require students to engage more deeply with mathematics, explain and justify their mathematical thinking, and communicate their thinking to one another (NGA, 2010). Thus, the documents that are intended to guide teachers' practice have embraced reforms to the traditional mathematics curriculum.

This more expansive view of algebra instruction does not, however, sacrifice the goals of skill mastery and computational fluency. The report of the National Mathematics Panel (2008) integrates both traditional and reform views by recommending that algebra curriculum should work to jointly foster conceptual understanding, procedural fluency, and problem solving. The Institute of Educational Sciences (IES) recommends instructional strategies that improve algebra knowledge through improving procedural knowledge, conceptual knowledge, and procedural flexibility (Star, et al., 2015). Thus, skill efficiency and conceptual understanding are seen as twin goals in modern algebra instruction (Hiebert, 2003; Star & Rittle-Johnson, 2009).

Policy-makers' instructional guidance to teachers in this regard has been shaped, at least in part, by the research base on student learning more broadly and algebra learning in particular. For example, in reviewing experimental and quasi-experimental studies across grades and content areas, IES recommends that, across grade levels and content domains, teachers alternate worked examples in which students examine previously solved problems, with independent work on similar problems (Pashler, et al., 2007). Such a practice has been shown to improve student understanding in mathematics more broadly. Similarly, IES found moderate evidence, mostly from experimental studies in laboratory settings, that students benefit when teachers connect abstract and concrete representations of topics. There is also strong evidence that students' academic performance is improved by instruction that encourages them to build their own explanations of concepts and ideas.

Recent research from educational psychology has identified instructional practices in algebra in particular that can support student learning (for a summary, see Star, et al., 2015). For example, having students analyze solved problems using algebraic reasoning improved students' conceptual and procedural knowledge in algebra (Booth, Lange, Koedinger, & Newton, 2013). Other research finds promise in comparing and contrasting multiple solution methods to problems (see for example, Lynch & Star, 2014b and Rittle-Johnson & Star, 2009) as a means to improve students' procedural flexibility. In addition, active and explicit connections across representational forms (such as graphs, equations, and tables) has potential for improving students' understanding in mathematics (Knuth, 2000; Star, et al., 2015).

How and Whether Algebra Instruction Has Evolved to Meet Reformers' Visions

Despite a growing understanding of instructional strategies that may be beneficial for student learning, there is scant evidence illuminating the degree to which the hoped-for practices have penetrated algebra classrooms, particularly at scale. Classroom-based research in algebra has explored why the subject is challenging for students (Kieran, 2007; Star & Rittle-Johnson, 2009), teachers' knowledge of algebraic concepts (Doerr, 2004; McCrory, Floden, Ferrini-Mundy, Reckase, & Senk, 2012), and how meaning-making practices *might* be translated into an algebra context in individual schools and classrooms (e.g. Boaler & Staples, 2008; Lynch & Star, 2014b; Raymond & Leinenbach, 2000). The last strand is particularly relevant to this study, in that it provides evidence, usually at small scale, of the nature of the instruction we may expect or hope to see in algebra classrooms. Yet research seldom looks broadly at whether this particular vision of instruction has taken root. The little research that does explore this question in large samples does so by investigating curriculum in use (e.g. Porter & Smithson, 2001), or surveying teachers' perceptions of their instruction (e.g. Banilower, et al., 2013).

One way in which instruction may have evolved to incorporate the instructional vision outlined above is through the wide-spread adoption of curricula that emphasize reform-oriented instructional practices. For example, after *A Nation at Risk* (Gardner, 1983) prompted the National Council of Teachers of Mathematics (NCTM) to develop curriculum standards, the National Science Foundation funded the creation, adoption and evaluation of curricula aligned to the reform vision outlined above (Cohen & Hill, 2000). Such curricula frequently advocate for student-centered instruction grounded in meaning-making around mathematical concepts (Borasi & Fonzi, 2002), include open-ended exploration, opportunities for student to make meaning of the mathematics they

encounter, and mathematics situated in real-world contexts (Boaler, 2002; Stein, Remillard & Smith, 2007). In a survey of a national probability sample of mathematics and science teachers, Banilower and colleagues (2013) find that while the NSF-funded curricula have been at least modestly adopted in the past two decades at the elementary and middle school level (with 25% and 11% respectively of teachers who use published curricula reporting their use) they have made little inroads at the secondary level. In this survey, fewer than 1% of secondary teachers reported using NSF-funded curricula. Thus it is less clear the degree to which high school algebra curricula emphasize reform-oriented practices.

It is possible, however, that teachers are engaging in these reform-oriented teaching practices irrespective of the curriculum materials they are using. Indeed, Banilower and colleagues (2013) find high levels of teachers' self-reporting reform-oriented practices in their mathematics teaching, indicating that teachers believe that their practice has shifted. For example, the vast majority of teachers surveyed indicated that students should be provided with frequent opportunities to share their thinking and reasoning about the mathematics in most class sessions. In addition, the majority of teachers believe that students should be given the opportunity to investigate a mathematical idea before having it explained to them by the teacher. It is important to note that while the authors find that teachers across grades report using reform-oriented practices, these practices are also more frequently reported by middle and elementary school teachers than by high school teachers. In another study, Lynch and Star (2014a) interviewed and surveyed middle and high school algebra teachers participating in a professional development study around teaching multiple strategies. Not surprisingly, the

researchers found that the majority of sample teachers believed that teaching students multiple strategies was a valuable endeavor and nearly 20% of teachers reported engaging in the strategy daily.

In contrast to teachers' self-reports, observational studies that have analyzed mathematics instruction present a mixed picture of the degree to which reformers' vision has permeated classrooms. On the one hand, some research has found that the instructional format of mathematics lessons remain, by and large, traditional. I define format of instruction here as overall lesson structures and characteristics such as the degree to which the instruction is teacher-led, the amount of the mathematical work done by the students, and the degree to which students engage in exploration or investigation of mathematical concepts. For example, descriptive analysis of the eighth grade TIMSS data finds that mathematics instruction in the U. S. continued, at least through the mid-1990s, to focus on procedures and definitions, with most lessons following traditional instructional formats in which the teacher directed the presentation of the content and students practiced previously-learned procedures (see for example Hiebert, et al., 2005; Jacobs, et al., 2006; Stigler & Hiebert, 1999). While this research did not focus specifically on algebra, much algebraic content was covered in these eighth grade lessons.

While the format of instruction appears to be relatively consistent over time, it is less certain whether particular instructional features—the more fine-grained instructional practices that teachers engage in—have made inroads into classrooms. An analysis of contemporary eighth and ninth grade algebra lessons using MET study data scored on project observation protocols finds little evidence of instructional features that might support conceptual understanding (such as providing mathematical explanations or

developing generalizations) or those that allow students to engage with mathematical content in cognitively activating ways (Litke, 2014). However, it may be that the instruments used to measure instructional practice in the MET study were not attuned to specific instructional features prevalent in algebra.

In contrast, an algebra-focused analysis of MET video (Chapter 2 of this dissertation) suggested the presence of several promising instructional features in MET video. Specifically, teachers engaged in multiple ways to leverage different types of connections (between representations, between mathematical topics, and between abstract and concrete algebraic ideas) that provided students the opportunity to develop deeper understanding of the mathematics. In addition, when teaching procedures, teachers engaged in activities such as making sense of procedures, supporting students' procedural flexibility, and attending to the organization of their presentation of the procedures. It may be that these features hold promise in building and developing student understanding in algebra. However, there is not yet evidence on the degree to which these features are prevalent in a large sample of lessons and if so, at what level of quality they appear.

Investigating the Nature and Quality of Contemporary Algebra Instruction

Thus it is an open question whether and to what degree particular instructional features have permeated algebra classrooms. The existing research on the nature of algebra instruction in typical classrooms tends to occur as case studies in one or just a handful of classrooms, and we know little about algebra teaching more broadly (Kieran, 2007). For example, Chazan (2000) used his own teaching of secondary algebra to analyze student engagement with content, the nature of the algebra curriculum, and patterns of discourse in his classroom. A descriptive analysis of a large sample of algebra

classrooms would add to our broader understanding of instructional practice in algebra and allow for future empirical investigations of whether (and which) instructional features hold promise for improving students' algebraic understanding. Similar work in eighth grade mathematics (e.g., Hiebert, et al., 2005) has proven influential in policymakers' thinking about how to improve teaching quality in middle school.

Methods

To answer my research questions, I take a mixed methods approach. I conducted descriptive analyses using an algebra-specific observational instrument in order to identify the prevalence of instructional formats and features across levels of quality. To describe the instruction in the sample in more detail, I complement these results with qualitative cases that depict instructional practices in algebra more deeply. Below, I first describe the data and sample used in this study. I next describe the classroom observation instrument used in this analysis, discuss the scoring protocol, and provide information regarding its reliability. Finally, I present my analytic strategy.

Data

Data for this study comes from the Measures of Effective Teaching (MET) Project. The project partnered with approximately 3,000 teachers across six urban districts: Charlotte-Mecklenburg, NC, Dallas, TX, Denver, CO, Hillsborough County, FL, Memphis, TN, and New York City, NY. Teachers on the project contributed videos of their lessons, which were then scored by MET project raters on a number of classroom observation instruments, both general and subject-specific. The project also collected teacher demographic information, administered assessments of teachers' content knowledge for teaching and student's mathematical knowledge (e.g. SAT-9), and

retrieved student achievement and demographic data from the partner districts. In addition, the project administered student, teacher, and administrator surveys. The overall goal of the MET project was to determine fair and reliable methods for measuring effective teaching (Kane & Staiger, 2012).¹⁹

Data Sources

Classroom video. This study relies heavily on the video-recorded classroom lessons from ninth grade teachers. Video-recorded classroom lessons allow for a deep analysis of the features of classroom instruction. Video also permits researchers to slow down and re-watch classroom moments that may be missed in live observation (Hamre, et al., 2007; LMTP, 2011). MET classroom lessons were video-recorded using an unmanned panoramic digital video camera operated remotely by teachers or school personnel (Bill & Melinda Gates Foundation, 2010a). While teachers selected the lessons to be recorded, the project requested that at least two of the lessons be focused on core topics in the teacher's subject area.²⁰ Teachers uploaded their videos to a secure website, where they identified the topic of the lesson and had the option to upload supporting materials.

Content Knowledge for Teaching: Algebra assessment. The MET project administered an assessment of pedagogical content knowledge to all teachers in the study. Those who taught ninth grade algebra took a specially-designed Content Knowledge for Teaching Algebra assessment (CKT) developed by researchers at the Educational Testing Service (ETS) and the University of Michigan (Bill & Melinda Gates Foundation, 2010b;

¹⁹ For more information on the design and components of the MET Project, see www.metproject.org.

²⁰ Analysis by Ho and Kane (2013) of a different subset of MET video data found that teacher-chosen video were rated higher on average than non-teacher chosen video, but the relative ranking of teachers did not change regardless of whether teachers chose their own video for scoring.

Gitomer, Phelps, et al., 2014). The CKT assessment was administered in early 2011 and focused on the specialized knowledge that might enable teachers to teach algebra content to students.²¹ The Algebra CKT exam included 37 selected response items (two of which were excluded for poor performance) and had an overall reliability of 0.77 (Gitomer, Phelps, et al., 2014).²²

Sample

For this study, I focused on the available ninth grade mathematics classroom videos from year one of the MET study. In particular, I focused on the subsample of ninth grade algebra lessons (233 teachers across five districts).²³ I excluded teachers with no scores on the Content Knowledge for Teaching (CKT) Algebra assessment administered by MET (92 teachers) and teachers who did not consent to make their video available to researchers (60 additional teachers), as both criteria were important to my sampling and analytic strategy (see below). My final analytic sample consisted of 81 ninth grade math teachers from 49 schools spread across five of the partner districts. Teachers in the final sample have between one and five viewable videos each, with most teachers having four videos, yielding a total of 292 video-recorded algebra lessons.

The analytic sample is largely similar in demographic composition to the full ninth grade sample, with some notable differences. The analytic sample is comprised of more white teachers (65% compared to 53%), fewer Black teachers (22% compared to 28%), and fewer teachers who are not White, Black or Hispanic (1% as compared to 6%).

²¹ For more information on the CKT assessments developed for MET see Gitomer, Phelps, et al. (2014).

²² Although this reliability is lower than traditional tests of teacher knowledge, it is important to note that the MET CKT assessment was shorter in length, and researchers calculated that doubling the length of the test would yield reliabilities of between 0.82 and 0.91 (Gitomer, Phelps, et al., 2014).

²³ While the larger MET study includes six partner districts, one district did not include any ninth grade math teachers as project participants; therefore this study is restricted to five districts. See Appendix A for a breakdown of district representation in the full sample, the ninth grade sample, and the analytic sample.

Analytic sample teachers for whom there is information on experience and education have been teaching for slightly fewer years on average (6 years compared to 7.5 years) but are more likely to have a Master's Degree or higher (32% compared to 28%). Comparing the sample of ninth grade teachers with CKT Scores (n=141) to those with CKT scores and viewable video (n=81), there are no discernable differences in mean, median or range of these scores. For a more detailed comparison of the analytic sample to the full ninth grade sample and the ninth grade sample with CKT scores, see Table 1.

As scoring all 292 videos was not feasible given the resources available to this project, I constructed a random sample that would be sufficiently large for the descriptive analysis I conducted and exceeds recommendations for qualitative research (Guest, Bunce, & Johnson, 2006). From the sample of 81 teachers, I constructed a subsample for scoring, randomly selecting 30 teachers proportionate to district representation in the full sample. This strategy yielded a total of 108 lessons (665 segments of instruction) and, because of random sampling, is representative of the larger sample of algebra lessons from which it was drawn.

Scoring Video Data

To describe the nature and quality of instructional practices in algebra lessons, I utilized the Quality of Instructional Practice in Algebra (QIPA), an algebra-specific observational instrument. Lessons were scored in 7.5-minute segments and the instrument includes two domains with four codes each at the segment level that focus on specific dimensions of instructional practice. In addition, the instrument includes two whole-lesson codes that focus on the overall format of instruction. In Tables 2 and 3, I define each of the domains and dimensions used in this analysis.

Within the segment-level codes, the first domain identifies particular features of instruction that teachers engage in *when teaching procedures* (see Table 2). I define procedures here as the instructions for completing a mathematical algorithm or process. The teaching of procedures occurred when the teacher demonstrated new procedures for the class, reviewed previously learned procedures, or engaged in a description of a procedure in the context of solving problems. Procedures are important as they are not only prevalent in algebra, but they are often more complex than those in arithmetic (Star, 2005; Star & Rittle-Johnson, 2009). As such, the instrument contains three codes to capture the quality of the instruction on procedures in a lesson. The first code (*Making Sense of Procedures*) captures the degree and depth with which teachers and students make sense of procedures, either by attending to meaning of the individual steps of the procedure, the solution generated by the procedure, or to the procedure as a whole. The next code (*Supporting Procedural Flexibility*) focuses on the degree to which teachers present algebraic procedures to students in ways that afford the opportunity to develop procedural flexibility. For example, this code captures such practices as noting multiple pathways through a procedure, noting when a particular procedure is appropriate or what cues the selection of that particular procedure, and comparing multiple procedures for their affordances or limitations. The third code in this domain (*Organization in the Presentation of Procedures*) indicates how complete, detailed, correct, and organized the teacher (or students') presentation of content is when describing or outlining a procedure.

The second domain, Leveraging Connections, is intended to capture the ways in which teachers made connections between and within mathematical ideas. It consists of three dimensions (see Table 3). The first code in this domain (*Connecting Across*

Representations) captures the nature of the connections teachers and students make between and across representational forms in algebra—specifically between graphs, tables, equations, and problem contexts. Connections across representations are an important way to develop meaning in algebra (Knuth, 2000) and this practice has been encouraged in recent decades of mathematics reform. The second code (*Situating the Mathematics*) looks at the connections teachers and student make across aspects of the algebra curriculum, to related topics, or to the broader domain of mathematics. Teachers do this in a variety of ways, including making connections between a current topic to prior or future content in algebra, using the architecture of the lesson to connect mathematical ideas together, or to connect what students are learning to the broader domain of mathematics. Finally, the third code in this domain (*Making Connections Between Concrete and Abstract Ideas in Algebra*) is intended to capture the degree to which teachers and students leverage concrete examples, representations, or ideas to develop understanding of abstract concepts, formulas, notation or definitions. Teachers do this, for example, by attending to what the components of a mathematical formula represent, clarifying mathematical definitions or abstract algebraic concepts with concrete examples (or non-examples), or explicitly relating an abstract concept to its analogous concrete idea.

In addition to these segment-level instructional features, the instrument also contains two codes that are scored at the lesson-level and capture instructional formats used in the lesson as a whole. The first lesson-level code (*Inquiry/Exploration*) is designed to capture the degree to which students were asked to do significant mathematical work involving mathematical investigation or discovery. A second code

(*Teacher-Led Instruction*) is intended to capture the degree to which the teacher directed the mathematical content and processes in the lesson and the degree to which students engaged directly with the mathematical content through such activities as independent practice or mathematical contributions. These codes are not necessarily in opposition to one another, as they focus on duration of these instructional formats. Thus, it is possible that a lesson contains times in which students engage in inquiry and exploration and times in which the teacher heavily directs the content. In addition, there are other lesson formats (e.g. completing test review worksheets in stations) that may not fit into either of these two categories. As such, these codes are not intended to be exhaustive of all possible lesson formats, but rather to capture the degree to which two particular formats noted in the literature are present in algebra lessons.

Lessons were scored using a rubric that captured the frequency and depth of each of the instructional features. Consistent with other observational instruments (e.g. Hamre, et al., 2007; Hill, et al., 2008; Hill, et al., 2014; LMT, 2011), lessons were scored at the segment level. Like Hill and colleagues (2014), instruction was scored in 7.5-minute segments, as these were short enough to keep track of mathematical practices and events, but sufficiently long to allow efficiency in scoring and to view instructional interactions in the broader context of the lesson. Raters scored each 7.5-minute lesson segment on each of the six codes capturing the Teaching Procedures and Leveraging Connections domains outlined above, using scores of Low (1) to High (5). A score of Low (1) indicates that the practice did not occur or that the teacher engaged in the practice but it was mathematically incorrect.²⁴ In contrast, a score of High (5) indicates that the

²⁴ Because a score of Low indicates an absence of a particular feature, I wanted to distinguish segments in which no procedures were taught (and thus scored Low on the codes in the teaching procedures domain by

particular instructional feature characterized the segment either by sustained focus on that particular feature or through an extended and well-developed instance of that particular feature. A score of Mid (3) was anchored to indicate “modal practice.” For each code, this indicates that a teacher engaged in the instructional feature more than briefly, but that the instructional feature neither characterized nor was the main focus of the segment. For example, a segment in which a teacher engaged in a particular instructional feature by stating a connection or relationship but without giving much further elaboration or discussion would score a Mid. Segments were scored Low/Mid (2) if the instructional feature was briefly present—even in passing. Finally segments were scored Mid/High (4) if the instructional feature was more developed than a Mid score would indicate, but did not reach the threshold for a High score. Whole-lesson codes (*Inquiry/Exploration* and *Teacher-Led Instruction*) were scored in a similar manner. After watching each lesson, raters gave the whole lesson a score ranging from Low (1) to High (5), indicating the degree to which instruction was teacher led or the lesson featured inquiry or exploration. For the complete scoring guide and instrument, see Appendix B.

An investigation of the instrument revealed reasonable measurement properties. (for a detailed description of the instrument validation and discussion of its validity, see Chapter 3). At the segment level, observers utilized all score points on each code, indicating a good fit of the instrument to algebra instruction. Among the 20% of lessons double-scored by a second rater, inter-rater agreement ranges between 0.36 and 0.65.

Correlating scores on the algebra-specific codes with other measures of teacher and

definition) and those in which procedures were taught but the teacher did not engage in the practices outlined in the codes. As a result, each segment was also scored for whether instruction on procedures was present or not. For the purposes of analysis, this allowed me to discuss the quality of instruction on procedures when procedures are taught.

teaching quality, I find that teacher's content knowledge is correlated with instructional format. Specifically, teachers' CKT score is positively correlated with the whole-lesson code of *Inquiry/Exploration* (0.6113, $p < 0.01$), indicating that teachers with higher CKT scores were associated with more student-centered instruction. The teacher-led instruction code was negatively correlated with CKT score (-0.4465, $p < 0.05$), indicating that teachers with higher CKT scores scored lower on average on this particular code, meaning that instruction tended to be less teacher-driven and more student-centered. In contrast, there are few statistically significant correlations between teachers' CKT score and the instructional features codes. There is however, positive correlation between teachers' CKT and the *Connecting Between Concrete and Abstract Ideas in Algebra* code (0.5182, $p < 0.05$). Interestingly, there is a negative correlation between teachers' CKT score and *Organization in the Presentation of Procedures* (-0.4351, $p < 0.05$). This may indicate that the CKT exam is not designed to measure this particular instructional feature.

Analytic Strategy

This analysis seeks to estimate the frequency and quality of key instructional formats and features, and to convey nuanced descriptions of these features in real classroom contexts. To do this, I first scored the 108 video-recorded lessons from this sample of 30 teachers with the QIPA. A member of the research team double-scored every 10th video. I used the scores on these algebra-specific codes to present descriptive statistics for each code to describe the prevalence, quality, and distribution of various practices in the sample.

Based on these results, I returned to the videos, developing case studies of both modal practice and high-quality instruction of these two domains of instructional practice

in algebra in order to better describe how these features of instruction manifest in real contexts (Yin, 2009). Doing so adds texture and nuance to the quantitative results and provides thick description of features of algebra instruction (Merriam, 2009). A case study approach allows for multiple sources of evidence to develop a broader understanding of a phenomenon (Stake, 1995). Descriptive cases help to illustrate “typical cases” (Yin, 2009)—what modal practice of a particular feature looks like—or what specific practices look like when they are present at a high level of quality (an “extreme case” as discussed by Yin, 2009).

To accomplish these goals, I selected particular score profiles of algebra-specific practices from the descriptive work above and purposively sampled lessons within those score profiles. I re-watched segment and lesson videos featuring these practices and developed illustrative cases of the algebra-specific dimensions of instructional practice. As recommended by Yin (2009), watching lessons for the purposes of case development was informed by a case protocol (see Appendix C). In developing the case descriptions, I relied primarily on lesson video, but supplemented this data with a combination of lesson and segment scores from the QIPA, scores on the MET observation rubrics, and teachers’ CKT scores. I watched each video using the case protocol and wrote analytic memos after re-watching each segment or video selected for this purpose, using concepts from the literature, as well as those that emerged from the lessons themselves, in order to describe the practices in context and explore patterns within and across lessons for each practice (Stake, 1995; Yin, 2009).

To select a lesson for the typical case, I followed a two-stage process as recommended by Yin (2009), using the quantitative data to narrow the pool of possible

lessons and then using a second screening stage to select the case. I first relied the descriptive results above to identify lessons that contained “average” scores as the lesson-level relative to other lessons in the sample. Specifically, I examined lessons in the third quintile of quality on the codes in the Teaching Procedures and Leveraging Connections domains at the lesson level. This yielded a group of 36 lessons, 15 of which scored in the third quintile on the Teaching Procedures domain, 16 of which scored in the third quintile on the Leveraging Connections domain, and five of which scored in the third quintile on both domains.

I next reviewed the lesson summaries from all 36 lessons and coded them for common themes. In reading and coding the lesson summaries, as well as re-examining the segment- and lesson-level scores for these lessons, I developed a list of themes that characterized instruction in this group. I aggregated these themes into broad categories. I then re-read the lesson summaries for this sub-set of lessons with the goal of selecting lessons that illustrated these themes. I re-watched each of the selected lessons using the case protocol, identifying and transcribing vignettes from each lesson that both supported the themes and provided contrasting evidence. Finally, I reviewed the lesson summaries, segment scores, and case protocol notes for each lesson and selected a single lesson that typified instruction in this group.

To select a case of high quality instruction, I utilized a similar process. To examine as large a sample of lessons as possible in which these instructional features occurred at high levels of quality, I selected all lessons with at least one segment that scored High (5) on any of the codes in either domain. I hypothesized that if lessons contained at least one segment that scored High on any of the codes, that there likely

would be other elements of quality instruction as well. This process generated a list of 24 lessons, nine of which had at least one segment that scored High on one of the Teaching Procedures codes, eleven of which had at least one segment that scored High on one of the Leveraging Connections codes, and two of which had at least one segment that scored High on both domains.

I followed a similar two-stage process as above, examining the scores on each segment on all codes for each lesson and analyzing the lesson summaries for each lesson, generating a list of themes as well as noting places of divergence. I selected a subset of lessons that illustrated these themes and re-watched the lessons using the case protocol. I then reread the lesson summaries and case protocol notes to select the lesson I present below.

Findings: Frequency and Quality of Instructional Practices

I begin by describing the observed instructional formats across lessons in the data and next discuss the distribution of the quality of the individual instructional features across segments of instruction.

Instructional Format

In Table 10, I display scores from 108 lessons on the whole lesson codes: *Teacher-Led Instruction* and *Inquiry/Exploration*. The majority of lessons in this sample contained significant amounts of teacher-directed instruction and, while there were often some opportunities for students to engage in mathematical work, this was most often done in limited segments of student practice. For example, among algebra lessons in the sample, 54% of lessons (n=59 lessons) scored Mid on the whole-lesson code *Teacher-Led Instruction*. A score of Mid on this code indicates that instruction was largely

teacher-directed but did contain some time for independent or group student practice of the mathematics. In these lessons, there was some evidence that students participated to some degree through mathematical contributions such as putting solutions to problems on the board, offering a description of steps in a mathematical procedure, or asking questions about the mathematics, yet in these lessons, instruction was largely teacher-directed. Furthermore, 33% of lessons scored above a Mid on this code (33 lessons scored Mid/High and two lessons scored High), indicating that in approximately a third of the lessons there was minimal opportunity for students to engage in doing mathematics independently or to engage in productive mathematical contributions. In some lessons, this included the teacher directing the content of the lesson for its duration—e.g. working through all of the mathematical examples, lecturing to students about mathematical formulas or properties, and instructing students to copy pre-written notes, thus offering little-to-no opportunity for students to solve problems or otherwise engage with the mathematical content. This format matches much of what other researchers have seen in other historical and large-scale analyses of mathematics instruction (e.g. Cuban, 1993; Hiebert et al., 2005; Hill, et al., 2014; Jacobs, et. al., 2006). While only two lessons scored Low on this code (indicating that instruction was largely student-centered with minimal direction from the teacher), 11% of lessons (n=12 lessons) included significant student independent mathematical work and less time with the teacher directing the content.

Examining scores on the *Inquiry/Exploration* code, I find that the vast majority of lessons (86%) included no student exploration or inquiry into algebraic concepts. There were no lessons that featured inquiry or exploration for the majority of the lesson. In 10%

of lessons, raters did note some student exploration, inquiry, or work on cognitively demanding tasks, but this work occurred for only part of the lesson. These results indicate that the student-centered instruction hoped for by the mathematics education and policy communities is largely absent in these lessons. In the majority of lessons, instruction is largely teacher-led, includes minimal student mathematical contributions, and seldom contains extended student exploration or mathematical inquiry. However, a focus solely on whole-lesson format may obscure variability in more granular features of algebra classrooms. I turn next to the distribution of scores on the segment-level codes addressing the features of instruction.

Teaching Procedures

As expected, the majority of lessons (95%) included at least one segment in which procedures were taught (n=102 lessons). In addition, teaching procedures occurred in 65% of all segments of instruction (431 segments). Each segment in which procedures were taught was scored on each of the three codes in the domain (*Making Sense of Procedures*, *Supporting Procedural Flexibility*, and *Organization in the Presentation of Procedures*). In what follows, I describe the distribution of segment-level scores on each code and then discuss the overall quality of teaching procedures at both the segment and lesson-level. Figure 6 shows the distribution of scores across codes. I discuss each code below.

Making sense of procedures. It is clear from Figure 6 that this feature of instruction was not a major feature of segments in the sample. Approximately two thirds of segments in which procedures were taught scored Low on this code (65%), indicating that procedures were presented with no attention to meaning or sense-making. Only one

segment of instruction scored High on this code, and fewer than 4% of segments scored above Mid. However, cursory references to the meaning of procedures was more common. In close to 21% of segments, teachers engaged in this practice in brief, isolated instances (e.g. a brief statement in passing about the meaning of the solution to a procedure or why a particular step works mathematically) scoring Low/Mid (2), and an additional 11% of segments scored Mid (3), indicating that in these segments, teachers and students made sense of procedures more than briefly, attending to meaning over several brief instances within a segment or for one, more sustained instance. For example, in a lesson on simplifying rational expressions, the teacher gave the students the expression $\frac{x}{\sqrt{2}}$ and introduced the procedure for rationalizing the denominator when there is a radical expression.²⁵ She explained to students that they could not have a radical in the denominator of the expression and needed to multiply the expression by $\frac{\sqrt{2}}{\sqrt{2}}$, which she demonstrated on the board. Although this part of the procedure was presented simply as a series of steps with no attention to meaning, she next discussed with the students why the procedure worked as intended, reinforcing the mathematical concept that multiplying an expression by one does not change its value. While this explanation was not sustained nor given in depth, it was developed more than briefly. Overall, however, it appears that this practice is not prevalent in the sample and, when it does occur, occurs briefly and is not the salient focus of instruction in the segment.

²⁵ For the purposes of anonymity and to preserve confidentiality, all teachers discussed are identified as female, regardless of their actual gender. In addition, while I endeavor to portray substantive elements of the instruction described in this paper as they occurred, details of specific mathematical examples/classroom interactions may be slightly altered so as to render descriptions of classroom moments unidentifiable.

Supporting procedural flexibility. Looking across segments at the degree to which teachers supported students in the opportunity to develop procedural flexibility, I find this practice to be slightly more prevalent than *Making Sense of Procedures*, although it too usually occurs briefly or superficially (see Figure 6). Although 47% of segments scored Low, indicating that procedures were taught with no attention to elements of flexibility, 27% of segments scored Low/Mid (27%) or Mid (16%) on this code. This indicates that teachers frequently attended to flexibility either only briefly or somewhat more than briefly, but that this instructional feature did not characterize the segment. For example, a teacher might mention to students that there is more than one way to solve a given problem type but not elaborate or demonstrate the multiple methods. Encouragingly, close to 10% of segments did feature scores of Mid/High (4) or High (5) on this code, indicating explicit attention to the elements of flexibility in a sustained way. Here, teachers might note two or more pathways through a procedure and demonstrate both, showing students how different decisions are made in each solution pathway and emphasizing that each pathway arrives at the same solution.

Organization in the presentation of procedures. It is clear from Figure 6 that in this sample, the teaching of procedures was largely well-organized and clear. Of the segments in which procedures were taught, approximately 68% scored Mid (3), indicating that the instruction was free of mathematical errors, reasonably clear and well-organized, and mostly complete. A smaller percentage of segments scored Mid/High (8%) or High (2%), indicating that the presentation of procedures included exceptionally systematic organization or thorough and explicit detail. For example, a teacher might have students write the steps of a procedure next to the steps of a mathematical example,

using arrows to draw correspondences between the written steps and where they appear in the mathematics. In 20% of segments, raters assigned scores of Low/Mid (2), indicating the presentation of procedures were not wholly complete or clear. These segments included vague descriptions of steps of procedures or confusing explanations of algorithms. Very few segments (2%) scored Low (1). A score of Low indicates that the instruction on procedures was incorrect or full of mathematical errors. Thus, teachers largely presented mathematically correct procedures with reasonable clarity. In some cases, however these presentations included thorough and careful, systematic organization while in other cases, the presentation was less clear and detailed.

Lesson-level quality of instruction on procedures. Simply looking at the percentages of segments of each code may obscure important elements of quality. For example, it is reasonable that a teacher may make a pedagogical choice to attend to procedural flexibility when introducing a new procedure by comparing it to previously learned methods, but that in the remaining segments of instruction, she may not choose to do so due to different instructional goals. To address this and to develop a more holistic picture of the quality of instruction on procedures, I next present results at the lesson-level. Rather than construct lesson-level scores by averaging scores across segments (Hill, et al., 2014; Kane & Staiger, 2012), I look instead at the percentage of lessons with at least one segment at each score point. I do this to try to capture the degree to which these practices are occurring *at some point* in the lesson, rather than present a numerical “average” level of quality in a given lesson.

The results of these calculations provides a more hopeful picture of instruction on procedures. In Table 11, I present the percentage of lessons with at least one segment that

scored at each successively higher level of quality. Looking at Row 1 of Table 5, it is clear that the majority of lessons contain at least one segment that scored above Low on all three codes in the Teaching Procedures domain. Specifically, in lessons where procedures were taught, 69% of lessons included at least one segment with scores above Low on *Making Sense of Procedures*, 86% of lessons including at least one segment with scores above Low on *Supporting Procedural Flexibility*, and nearly all lessons included at least one segment that scored above Low on *Organization in the Presentation of Procedures*. This indicates that all three instructional features occur—at least to some extent—in the majority of lessons in the sample. Even more encouraging, 37% of lessons contained at least one segment scoring Mid or higher on *Making Sense of Procedures*. This indicates that despite there being few individual segments scoring Mid or above on this code, the instances of these practices are spread across more lessons than the segment-level prevalence might indicate. Similarly, over half of lessons contained at least one segment scoring Mid or above on the *Supporting Procedural Flexibility* dimension, and approximately one fourth of lessons contained at least one segment scoring above Mid on this code. Taken together, it appears that these practices are present in algebra lessons in this sample, at least to a modest degree.

Leveraging Connections

Next, I present results for the frequency and quality of the practices in the Leveraging Connections domain. All segments (n=665) were scored on the codes in this domain regardless of whether procedures were taught. The practices captured by this domain can occur in the context of teaching procedures, but can also occur in the context of student exploration, describing concepts, introducing mathematical definitions and

other elements of mathematics lessons. In what follows, I describe the distribution of segment-level scores on each code in the Leveraging Connections domain and then discuss the overall quality of leveraging connections at both the segment and lesson-level.

Connections across representations. Much has been made in the mathematics education literature of the promise and importance of making connections across representational forms in mathematics such as graphs, tables, equations, and problem contexts (see for example Knuth, 2000). Connecting across representations features prominently in many reform-oriented curricula and in recent and current standards. Despite this, this feature of instruction did not appear to be prevalent in the sample. As is clear in Figure 6, a large majority of segments (71%) scored Low on this code. A score of Low indicates that only one representation is present, multiple representations are present but no connection is made between them, or the connections made between representations are incorrect. It is important to note that some segments may have scored Low because the particular content in the segment did not lend itself to connecting across representations (e.g. using the rules of exponents to simplify rational expressions), but other times segments scored low because the teacher did not actively make connections across representations (e.g. having a graph and a table displayed but doing nothing to connect these two representations). In 7% of segments, teachers made brief connections between mathematical representations. For example, in discussing solving a linear equation, one teacher said, “you could also graph this problem to find the answer,” indicating the existence of and a brief connection to another representation, but did not elaborate further.

In some segments, teachers did engage in explicit work connecting algebraic content across representations. Thirteen percent of segments scored Mid on this code, meaning the teacher or students discussed a connection across representations more than briefly, but did not explore those connections in depth. For example, in a lesson on writing linear equations, a teacher connected the y-intercept from the graphical representation of a line to the y-intercept in the linear equation. She used the coordinates of the y-intercept to help determine the slope of the line algebraically. She then showed students how to find the slope using only the graph. While she pointed to the correspondences between the graph and the equation, she did not discuss or elaborate on the connection further. While 8% of lessons scored Mid/High (4), indicating some careful work making explicit connections across representations, only 1% of segments scored High on this dimension, demonstrating explicit, detailed and in depth connections across representations.

Situating the mathematics. As is clear in Figure 6, this instructional feature was the most rare in the sample. Intended to capture the connections that teachers and students make to other aspects of the algebra curriculum or to related mathematical topics, this code reflects the degree to which teachers motivate the current mathematical content within a broader context. Notably, 91% of segments scored Low on this code, indicating no explicit situating of the mathematics. Frequently, mathematical topics were presented as discrete and disconnected entities. For example, in one segment introducing trigonometric ratios, a teacher began by telling students they were learning trigonometry and defined sine as opposite divided by hypotenuse, giving no context or background for this new area of study nor connecting it in any way to other topics in the course or the

lesson. Fewer than 1% of segments scored High, meaning that very few segments featured instruction that explicitly attended to how mathematical topics or ideas are connected and elaborated on the nature of that connection. Furthermore, only 2% of segments scored Mid (3), in which teachers connected mathematics to prior or future content to some extent, but did not develop these connections. For example, in one lesson on graphing quadratic equations, the teacher took a moment to make connections to the students' prior experiences with a study of linear functions. She noted correspondences between the forms of the linear and quadratic equations, attending to their similarities and differences, but did not elaborate further.

Connections between concrete and abstract ideas in algebra. Given the abstract nature of algebra (Booth, 1988, Chazan & Yerulshlami, 2003; Rakes et al., 2010) and the difficulties this creates for students (Stacey & Chick, 2004), there may be promise in working to explicitly ground abstract algebraic content in concrete mathematical ideas and objects. For example, teachers might unpack an algebraic formula, attending to what the symbols in the formula represent or clarify an abstract algebraic concept using examples and non-examples. In Figure 6, I show the distribution of scores on this code. These types of connections were engaged in to some degree in approximately one quarter of segments in the sample, but similar to the other codes in this domain, were usually made briefly and without much elaboration. Here, 11% of segments scored Mid, indicating that in these segments, teachers more than briefly commented on connections, but these were not the focus of the segment. For example, in introducing slope-intercept form of a line ($y = mx + b$), a teacher might comment that the m and b are constants rather than variables, articulate that the m represents the slope and

the b represents the y-intercept, and proceed to label these components of the formula. In 9% of segments, these connections happened in brief, isolated instances. In only 1% of segments were these connections the major feature of the teacher-student work in the segment and in 74% of segments, this feature was completely absent.

Lesson-Level quality of leveraging connections. As above, I next examined the percentage of lessons with at least one segment at each score point on the codes in this domain. I do this acknowledging that teachers make multiple instructional decisions over the course of a lesson and that there may be good reason not to engage in some of the practices captured in this domain in any given 7.5 minutes of instruction. Focusing only at the segment level may obscure the prevalence of these practices across lessons, as it is unclear for example whether and to what degree segments are clustered within specific lessons or spread out among them.

In Table 12, I again aggregate these results to the lesson level, showing how frequently these instructional features occur within lessons in the sample. Each row presents the percentage of lessons with at least one segment at each of four levels of quality. Looking across codes in this domain, it is clear that there are some types of connections that teachers are more likely to engage in than others. For example, as indicated in Column 3 of Table 12, 71% of lessons contained at least one segment that scored above Low in *Connections Between Concrete and Abstract Ideas in Algebra*, indicating that in these lessons, the teacher made at least a brief connection between abstract algebraic concepts and concrete ideas or underpinnings. In contrast, Column 2 in Table 12 shows that only 39% of lessons featured at least one segment that scored above Low in *Situating the Mathematics*. This indicates that lessons far less frequently included

instruction in which teachers situated the mathematics under study in the broader mathematical context, even briefly. Encouragingly, close to half of lessons contained at least one segment that scored Mid or higher in *Connections Across Representations*, and 29% of lessons contained at least one segment that scored above Mid on this code. This indicates that this instructional feature has begun to make its way into classrooms at some level. Looking across the codes in this domain, it appears that these instructional features do occur to varying extent across lessons. But it is also clear that only a small percentage of lessons (ranging from 2% to 7% across codes) contain at least one segment of instruction in which these practices occur at high levels of quality. This implies that there may be productive work to be done to build teachers' capacity to deepen these practices.

Illustrations of Instructional Practices

In order to better understand what existing practice around these instructional features looks like—both at “typical” and high levels of quality—I next present two illustrative cases of lessons from the sample. I first describe a “modal” lesson—one that was typical of the quality of instructional features presented above. After this, I present a lesson illustrating what the practices outlined above look like when practiced together and at high levels of quality. While lessons with this second profile were rare in the sample, if the goal is instructional improvement, it is thus useful to illustrate a positive case to illustrate what engaging in these features looks like at higher levels of quality. Such “extreme” cases (Yin, 2009) may also help to bridge the gap between where “typical” instruction lies and where we may hope it to be.

Typical Instruction: The Case of Ms. Close²⁶

To illustrate typical instruction, I first selected those lessons that scored in the third quintile on the Teaching Procedures and Leveraging Connections domains from the QIPA. I relied on segment scores and lesson summaries to develop themes that categorized “typical” instruction in this sample. Lessons in this group featured a heavy focus on teaching procedures and included Low scores on many of the dimensions of the instrument across segments. This was sometimes due to content that did not lend itself to particular features. But other times, these Low scores represented missed opportunities to engage in these practices. Yet alongside these missed opportunities, lessons in this group also featured instructional features measured by one or more codes in the QIPA instantiated at middling levels of quality. In addition, lessons also featured occasional moments of higher-quality instruction in one or more of the instructional features. For example, these lessons included a number of isolated scores of Mid or Low/Mid in various codes across multiple segments or featured a single segment in which a number of practices in the instrument were present at a Mid level, with the rest of the lesson scoring Low on most codes.

This profile is apparent in a lesson on multiplying two polynomials taught by a teacher I call “Ms. Close.” To start the lesson, Ms. Close told the students they would be learning “a combination of some old concepts that you’ve already learned how to do” and new procedures. Although she alluded here to a connection to prior content, she never specified the particular content to which she was referring, missing an opportunity to situate the mathematics of the lesson in prior content. Ms. Close next assigned the students two example problems: $3(x + 3)$ and $2x(x - 2x)$. She asked the students:

²⁶ Teachers’ names are pseudonyms.

- Teacher:* When you look at this practice problem right here, what do you see? What do you need to solve that? What do you need to know?
- Student:* How to add.
- Student:* You have to multiply it, like you distribute it.
- Teacher:* What is that called? Distributive...
- Student:* Property.
- Teacher:* Distributive property.

Here, the connection to prior content is implied—Ms. Close frequently mentioned the distributive property in the early part of the lesson—but not made explicit for students. This interaction is an example of another missed opportunity to situate the mathematics, as Ms. Close opened the door to make a connection to prior content but did not follow through and make the connection explicitly for students.

Yet in this same segment of the lesson, Ms. Close also briefly supported procedural flexibility, although she did not instantiate this practice in depth. After assigning students two more problems to complete and reviewing their solutions on the overhead projector (See Figure 7), Ms. Close reminded students they needed to “simplify” their answer by combining like terms. She briefly compared the solutions to the two example problems noting:

There are going to be times where you are going to multiply using the distributive property and you’re going to get $3x + 9$ in this case here—and it was simplified—there were no terms that were alike, there was nothing else we could do. That was our final answer. Then there may be a time when you’re distributing and there are going to be like terms. Now we could have done this in the middle—those are like terms in there—but I wanted you to see the whole distributive property method. And then we combined it at the end.

In this instructional moment, Ms. Close briefly referenced two possible ways to solve the second example—by distributing the $2x$ first and then combining like terms as she had demonstrated it or alternatively by first subtracting $x - 2x$ and then distributing. While there was certainly potential in this moment for

supporting procedural flexibility, this mention of multiple methods was done only briefly and was not elaborated upon.

In a later part of the lesson, Ms. Close introduced a new procedure for multiplying binomials, using the common mathematical acronym FOIL, saying:

What you just multiplied was a binomial by a binomial. So now there's another method – another way that you can do the same thing. So we're learning two different methods on how you can solve a binomial times a binomial.

Here, Ms. Close briefly worked to support procedural flexibility, indicating that there is an additional method to do similar types of problems. In this instruction that followed however, she did not capitalize on this opportunity. She then introduced this new technique, a common procedure for multiplying binomials with the acronym FOIL:

A lot of teachers refer to it as the FOIL method... It's an acronym to help you remember how to multiply these. Okay? So the FOIL method. The word FOIL is an acronym for First, Outer, Inner, and Last. And what we're going to do is multiply.

She gave students a new example and narrated the procedure of multiplying the first terms of the two binomials, the outside terms, the inside terms and then the last terms. Despite her introduction of FOIL as another method to achieve the same ends as distributing, she did not return to this idea and did nothing to connect these two methods nor to compare them. In fact, both methods rely heavily on the distributive property and are actually different variants of the same mathematical idea, a similarity that went unexplored.

Despite these multiple missed opportunities, the lesson also contained moments in which Ms. Close engaged at least briefly, and at times more than briefly, in a number of

the instructional features measured by the QIPA. For example, later in the lesson, Ms. Close engaged in supporting procedural flexibility when she demonstrated the procedure for multiplying a binomial by a trinomial. She presented students with the example $(x - 2)(x^2 + 3x - 5)$ and demonstrated what she called “the horizontal method”—where each term in the binomial is multiplied by each term in the trinomial and the resulting expression is simplified by combining like terms (See Figure 8, Panel 1). Next, she introduced a new method for multiplying the same two binomials, what Ms. Close called “the vertical method.” She called it the vertical method because the two expressions were written on top of one another. She compared this new method to the previous method, stating:

- Teacher:* This was a horizontal alignment. You remember you used to multiply back in third, fourth grade? And you multiplied and you did long multiplication? You multiplied straight down. You can do that with polynomials...
- Student:* If we already get the first method?
- Teacher:* I’m going to show you this and you can choose if you either want to get with this way or get with this way. Everybody processes information differently. Okay? So I’m not going to box you into just knowing it this way. If *you* get it this way. Even if the problem was written like this [points to vertical alignment]. If you saw it written like this. I took the same problem. Sometimes—I don’t know who writes the test. They make give you this in a vertical alignment or like this. You can change this to make it look like that, okay? Or you can just go ahead and multiply it. So let me show you how to do it this way.

Here, Ms. Close elaborated a bit more on the idea of there being two methods to multiply a binomial by a trinomial. She indicated the possibility of students choosing the method that worked best for them, reminding students that no matter how the problem was presented, they could manipulate it to utilize whichever method they preferred. She also briefly connected the vertical method to the U.S. standard algorithm for multiplication of

numbers, connecting the abstract idea of multiplying two polynomials to what students knew about multiplying numbers.

She next demonstrated this “vertical method,” which followed the same steps that multiplying a two-digit by three-digit number would (see Figure 8, Panel 2). In the vertical method, she first multiplied the -5 from the trinomial by the -2 from the binomial (as you would with the ones digits in the standard algorithm for multiplication); she next multiplied the $3x$ by -2 and the x^2 by -2 . She then placed a zero in the “ones place,” again as you would in the standard algorithm for multiplication and multiplied each term in the trinomial by the x from the binomial. In using the same example as she had used in the horizontal method, she was able to demonstrate that the two procedures resulted in the same answer. She briefly compared the processes, stating that the vertical method did the same steps as the horizontal method but in a different order. She concluded this segment of instruction by reminding students:

So you guys have a choice. If it starts out like this—see how I rewrote it—
If it starts out like this you can rewrite it like this. So you have a choice.
You can either do a horizontal alignment when you’re solving them or you
can do the vertical alignment. Either or.

In this example, Ms. Close engaged in explicit work around supporting procedural flexibility. She demonstrated two different methods using the same mathematical example, making explicit for students that it would be possible to use either method to multiply the two polynomials. While this practice was not engaged in at the highest level of quality, as the connections were not elaborated on in depth, it was more than briefly present in this portion of the lesson. To a lesser extent, Ms. Close also situated the mathematics, briefly comparing the vertical method to the procedure for multiplying multi-digit numbers (“You remember you used to multiply back in third, fourth grade?”

And you multiplied and you did long multiplication? You multiplied straight down. You can do that with polynomials”). Overall, this vignette illustrates the ways in which various instructional features appeared in Ms. Close’s lesson. While these features did not characterize instruction in the lesson, they appeared at least to some extent.

In summary, this lesson was characterized by multiple Low scores on a number of codes across multiple segments as well as some Mid scores on particular codes across multiple segments. It contained brief instances of supporting procedural flexibility and situating the mathematics at low to mid levels of quality. These practices were not elaborated upon, but Ms. Close made attempts to engage in them, at least superficially. While there was some focus on supporting procedural flexibility through using and comparing different procedures for the same type of problems, this was not done in depth. The lesson also contained missed opportunities to engage in these practices (and others). There were opportunities to support procedural flexibility in the earlier portions of the lesson that were not capitalized upon. Similarly, while Ms. Close mentioned mathematical properties that would allow her to make sense of the procedures she was presenting, she did not make use of these to explicitly do so.

Typical Instruction: Missed Opportunities and Glimmers of Promise

Ms. Close’s lesson was typical of instruction in this group in many ways. It featured both instances of instructional features present briefly or at a middling level of quality, as well as missed opportunities to engage in some of the instructional features at high levels of quality. One theme common among lessons in this category was that there were notable moments in which teachers’ instruction was primed to engage in the practices in the algebra instrument, but these moments were not taken up. This

manifested in multiple Low scores across codes and segments, despite the foundation being present for higher scores. Thus, lessons frequently contained what I have categorized as “missed opportunities” for engaging in these instructional features. For example, Ms. Close surfaced multiple ways to solve problems in her lesson, which could have been made more explicit. These methods might have been compared or correspondences between them elaborated upon, but they frequently were not. Thus, by missed opportunities, I mean that while teachers’ instruction frequently set up the opportunity to engage in particular practices, the teacher did not always take up these opportunities.

Teaching is a complex endeavor in which teachers must make multiple decisions about what and how to engage students in content over the course of a lesson. There may indeed be good reason *not* to engage in a particular instructional feature in a particular instructional moment. For example, a teacher aiming to develop proficiency in a particular procedure with students may not wish in a given segment or lesson to discuss multiple solution methods for a given problem. It may in fact be true that the teacher *does* engage in this practice in a subsequent lesson. Thus, I do not present instances of missed opportunities as evidence of pedagogical weakness or ineffectiveness. Rather, I view these missed opportunities as important because they provide insight into particular instructional moments in order to discuss how *and whether* instruction might benefit from engaging in the practices under discussion.

Alongside repeated missed opportunities, the lessons in this group also featured “glimmers” of the practices captured in the observational instrument. For example, some lessons in this group contained single 7.5-minute segments of instruction, which featured

multiple practices on the instrument instantiated at middling levels of quality, while the rest of the lesson featured no evidence of these practices. Alternatively, some lessons had glimmers of one particular instructional feature enacted briefly in multiple segments. For example, in Ms. Close's lesson, the segment in which she compared the horizontal to the vertical methods for multiplying binomials provided one such glimmer. Here, she engaged in supporting students' procedural flexibility with some explicit focus, although this focus was limited to this portion of the lesson. In summary, across lessons, in this category, the instructional features captured by the instrument were present to some degree, but were not engaged in at great depth or with much elaboration.

Instructional Features at High Quality: The Case of Ms. Rose

Although most lessons featured only brief instances of the practices discussed above, there were some lessons in which the teachers engaged in these instructional features at high levels of quality. In this section, I first present a case of a single lesson (taught by a teacher I call Ms. Rose) that is illustrative of lessons in this group. Next, I discuss general themes that emerged examining this group of lessons as a whole.

Ms. Rose's lesson focused on graphing linear inequalities—a topic common in the Algebra I curriculum and in this data set. This particular lesson was characterized by High scores on almost all the codes at some point during the lesson, with some segments scoring High on multiple codes simultaneously. In what follows, I present the overall arc of the lesson, including vignettes of instruction, to highlight some of the moments of high quality instruction in both teaching procedures and leveraging connections. I conclude with a summary of the characteristics of the lesson as a whole.

Ms. Rose actively made sense of the procedures being taught throughout the lesson. For example, the lesson began with students working individually to complete a warm-up problem in which they were asked to graph the inequalities $x \leq 2$ and $x > -3$ on a number line. The students worked for approximately three minutes before Ms. Rose had a student put the solution on the Smart Board, and had her explain it. She stopped the student numerous times to focus on the meaning behind the student's solution. For example, she noted that in the first example, the student drew a ray beginning with a closed circle on the value of $x = 2$ and covering all values on the number line less than two (See Figure 9, Panel 1). She then asked the class:

- Teacher:* Why is it important to use a line with the circle? What does this all represent? [Points at ray].
- Student:* It's continual.
- Teacher:* It's continual. Okay. It continues, going on. We have an infinite number of dots there alright...
- Student:* What x is. What x can be. ... It represents what x is and can be.
- Teacher:* Right. Exactly. All possible values of x , is another way to say that. That's great.

In this exchange, Ms. Rose worked to make sense of the procedure for graphing an inequality on a number line by attending to the meaning of what the ray represents in the solution, underscoring that the reason the solution is a ray (rather than a number or a circle for example) is that it is meant to represent all values of x that make the inequality true.

Ms. Rose's lesson was also characterized by multiple instructional features from the QIPA at mid to high levels of quality. For example, when teaching students the procedure for graphing linear inequalities on the coordinate grid, she engaged in making sense of the procedure, attending explicitly to the meaning behind the shaded region and what she called the "boundary line." During this same segment of instruction, she also to

a lesser degree situated the mathematics by connecting the new topic to the warm-up, made connections between an abstract mathematical idea (graphing inequalities in the coordinate grid) and a concrete representation (the number line), and connected across representations (symbolic to graphical). In this part of the lesson, she displayed for students a blank coordinate grid and the inequality $x \leq 2$, the same inequality from the warm-up. She copied and pasted the number line graph from the previous screen onto the new screen on her Smart Board and asked a student to place the number line solution in the appropriate place on the coordinate grid (See Figure 9, Panel 2). After a student did so, Ms. Rose addressed the class, again situating the new topic in the context of prior mathematics, explicitly connecting the two:

- Teacher:* How does this grid differ from what we've been working with in the Do Now?
- Student:* [Points to x- and y-axes] It goes horizontally and vertically.
- Teacher:* Our number line only goes in one dimension and this goes in two dimensions. So because of this being two-dimensional, how might our solution differ?
- Student:* There will be coordinates?
- Teacher:* There would be coordinates instead of just a single point, x-y-coordinate. That's good... is this the only place on the graph where we've drawn this that $x \leq 2$? [Points to copy of number line solution]
- Student:* No. We could move it up or down.
- Teacher:* We could move it up or down. Would anyone like to do this?

Here, Ms. Rose demonstrated to students the difference between the concrete number line representation and the coordinate grid representation, beginning to illuminate the ways in which the solution to the problem in two dimensions will differ from the solution in one dimension. In doing so, she also connected the coordinate grid to the concrete number line with which students appeared quite familiar. Next, Ms. Rose had a student come to the Smart Board and copy (clone) the ray, placing it on the coordinate grid beginning at x

= 2, but along a different y-value. Ms. Rose then had the student make multiple clones and continued to overlay these clones onto the graph (See Figure 9, Panel 3).

After the student finished, Ms. Rose asked:

- Teacher:* What's happening folks? What have we done? I don't just have a point or a line anymore do I? What have we got?... What's happening here?
- Student:* It all stays on two.
- Teacher:* It stays on two. So we have almost a line don't we? [Points to closed circles along $x = 2$]. As a matter of fact, if we put enough of these, what we call clones, the dots would actually make what?
- Student:* A vertical line
- Student:* A line.
- Teacher:* A line. Would this line be a solid line do you think?
- Student:* Yes.
- Teacher:* So we would have some sort of a solid line here. And a shaded... what we call a shaded region here.

Here, Ms. Rose made sense of the procedure for graphing a linear inequality by making meaning of the shaded region and the “solid” line. She also connected the new abstract concept of graphing linear inequalities to the more concrete tool of a number line. In doing so, she made connections between prior mathematical content and the algebraic content she is teaching, situating the mathematics quite explicitly. In addition, she was beginning to work toward connecting the symbolic inequality with its graphical representation, as it is the inequality that generates the shaded region in the graph.

In the second half of the lesson, Ms. Rose focused more deeply on the connections between the symbolic representation of linear inequalities and their corresponding graphs. She introduced a screen that had a generic linear inequality: $y < mx + b$, with buttons (called sliders) attached to the inequality symbol, the m , and the b . At the outset, m was set to equal zero, b was set to equal -1 and the inequality symbol was set to \geq . This yielded the inequality $y \geq -1$. Next to the symbolic form was a

graphical representation of the inequality. The software had these two representations linked such that changing the values of m , b , or the inequality symbol would display the corresponding change in the graph. The dynamic set-up of the sliders allowed Ms. Close to toggle between $<$, \leq , $>$, \geq , and $=$, altering the symbolic representation and seeing how that change manifested in the corresponding graph. Similarly, she could also change the value of m or b and the corresponding change would be displayed in the graphical representation.

After explaining to students what the sliders were, she turned students' attention to the inequality $y \geq -1$ and its corresponding graph and said:

- Teacher:* I want you to observe what happens as I hit the different sign, what changes in the graph. Because then you're going to have to come up with a rule for what is happening.²⁷
- Student:* I think it's the direction of the arrows.
- Teacher:* The direction of the arrows, you mean the shading?
- Student:* Yeah
- Teacher:* Potentially. Let's see what happens. We're on $y \geq -1$ now I'm just going to hit greater than instead of greater than or equal to. Okay? What changed?
- Student:* [Inaudible].
- Teacher:* Now I'm going to hit the equal sign can anyone going to guess what's going to happen now?
- Student:* It's going to be one line.
- Teacher:* One line. Any shading up or down? No shading at all. If we do the equal sign, we're actually just looking at the linear equation, which is what we've been working with so far. Now I'm going to hit less than. What happened there?
- Student:* Shading is now going down.
- Teacher:* The shading is now going down. What's going to happen when... I go from less than to less than or equal to, what's going to happen to the graph? What do you think?
- Student:* It's going to change direction.
- Teacher:* It's going to change direction? What do you think?
- Student:* I think it's going to be a solid line.

²⁷ Technically, the correct mathematical term here would be symbol so as to distinguish the greater or less than symbol from a positive or negative sign. While I acknowledge this imprecision, it is a common one in practice and Ms. Rose is clear that she is referring to the inequality symbol in this vignette by both her gestures and the set-up of the button on the display.

Teacher: It's going to be a solid line because I went from less than to less than or equal to. And she's correct. Now I'm going to quickly bounce back those buttons to the other way. And when we're done we need a rule. You're looking for the relationship between the type of line and the sign.

Ms. Rose used the sliders to develop the relationship between the inequality symbol and its manifestation on the graph. She linked the symbolic and graphical representations, showing students the connection between the inequality symbol and the boundary line, as well as the connection between the direction of the inequality and the location of the shaded region. She next formalized these connections in a highly structured and organized table that she used to guide students in the procedure for graphing linear inequalities. It is worth noting that in this segment, while the connections across representations were made explicitly and characterized the instruction, the uncovering of these relationships was highly teacher-directed.

In addition to the focus on connecting across representations in this segment, Ms. Rose also situated the mathematics once again, connecting the relationship between the symbolic and graphical forms of linear inequalities to what students already knew about linear equations. For example, she connected the m and b in the general form of the linear to the slope-intercept form of linear equations ($y = mx + b$). She went further with this connection, discussing it in the context of a new inequality displayed on the Smart Board: $y < -1$, identifying the slope and y-intercept for this particular example and discussing why a slope of zero yielded a horizontal boundary line.

On the surface, the instruction in this lesson might be considered typical: Ms. Rose did much of the cognitive work in the lesson and students, while engaged, were mathematically passive for large portions of the lesson. In fact, it was only at the end of

class that students were asked to do mathematics on their own, completing a worksheet on the material Mrs. Rose had just presented. Thus the instructional *format* was similar to that found in prior studies of U.S. mathematics classrooms (e.g. Hiebert, et al., 2005).

However, this lesson serves as an illustration of the ways in which particular instructional *features* might be instantiated at high levels of quality. For example, Ms. Rose frequently engaged in making connections across representations (specifically between graphical and symbolic representations) at high levels. At various points in the lesson, she made sense of the procedure for graphing linear inequalities, focusing on both the purpose of the procedure, the meaning behind individual steps and the meaning of the solution. In addition, there were a number of segments in the lesson in which she explicitly situated the new content she was teaching in the larger domain of mathematics, with multiple segments scoring High on this code. While the lesson did not feature a sustained focus on the other instructional practices highlighted in this paper, there was at least one segment that scored High on each feature.

Despite numerous instructional features at high levels of quality, this lesson—like most others in the sample—also featured some missed opportunities to engage in the particular instructional practices discussed in this paper. For example, in writing the symbolic inequality from a given graph, Ms. Rose told students that to find the slope they needed to pick two points and apply the slope formula. Specifically, she told them to pick the x- and y-intercepts. While there is nothing incorrect about this method, it is not the only method for finding the slope from a graph. Here might have been an opportunity to engage in instruction that supported procedural flexibility, but instead Ms. Rose emphasized one particular procedure as “the way” to go about writing the inequality. It is

certainly possible (and reasonable) that Ms. Rose may have had good reason to not engage in this practice at this particular instructional moment, but it was an opportunity not taken up nonetheless.

High Quality Lessons: Depth on a Single Dimension or Feature

I present the description and analysis of Ms. Rose's lesson above in order to describe a lesson in which multiple practices were enacted at high levels of quality as it is potentially useful to see how these practices may be instantiated in tandem. While Ms. Rose's lesson featured instruction that scored High across segments in multiple codes across both domains, this was unusual among the group of high quality lessons. In general, lessons with at least one segment that scored High on one of the codes in the Teaching Procedures domain did *not* have segments that scored High in the Leveraging Connections domain (and vice versa). As stated above, there were only two lessons with high scores on codes in both domains. This may indicate that while individual practices may be featured at high quality in lessons, they may be content- or context-specific. It is worth noting that while Ms. Rose's lesson content may have been a visible opportunity to engage in making connections across representations (which would arguably have been more difficult with the content Ms. Close was teaching for example), both case teachers had opportunities to engage deeply in one or more of the instructional features captured by the instrument. It may not be necessary (or desirable) for teachers to engage in multiple features simultaneously, rather instruction should focus on appropriate features in a given context. Indeed, many of the lessons in this subset went deep into one particular instructional feature throughout the lesson, with teachers engaging in that feature at high levels of quality, but did not include a focus on the other features.

Discussion & Conclusion

In this paper, I aimed to understand the nature and quality of instruction in contemporary algebra classrooms. I used the QIPA, an algebra-oriented classroom protocol, to score 108 lessons from 30 teachers randomly sampled proportionate to district representation in an effort to describe the frequency and quality of algebra-oriented practices in ninth grade algebra lessons from the MET project sample. Finally, I presented cases of algebra lessons to illustrate in greater detail how these features are instantiated (or are not instantiated) both in typical lessons and in high quality lessons in this sample.

Like others who have engaged in large scale studies of mathematical practice in U.S. classrooms (e.g. Hiebert, et al., 2005; Kane & Staiger, 2012), I find evidence that traditional formats of instruction persist. Instruction in this sample is largely teacher-directed, contains little evidence of student inquiry or exploration in high cognitively demanding mathematical tasks. In addition, I find that instruction on procedures is prevalent. In the vast majority of lessons, procedures are presented clearly and free of mathematical errors. There is some attention to supporting procedural flexibility in over half the lessons and many lessons feature brief attention to making sense of the procedures, at least for part of the lesson.

Yet none of the practices were engaged in at high levels with any frequency. Lessons featured high scores on *Supporting Procedural Flexibility* the most frequently, but even so only 8% of lessons contained at least one segment scoring High on this code. Some practices rarely appeared to be instantiated at high levels at all. Only 1% of lessons featured at least one segment that scored High on *Making Sense of Procedures* and 2% of

lessons featured at least one segment that scored High on *Situating the Mathematics*.

Thus it seems that while it is somewhat common for teachers to attend briefly to these practices during their instruction, these practices are not often engaged in at high levels of quality and do not characterize the instruction.

One notable finding is that some instructional features are more prevalent than others in the sample. For example, 90% of lessons contained a segment that scored Mid or higher in *Organization in the Presentation of Procedures*, while only 37% of lessons scored at that level in *Making Sense of Procedures*. This finding may not be surprising as making sense of the steps of a procedure, explaining the mathematical reason why a procedure holds, or making sense of the solution generated by a procedure may require more mathematical knowledge than simply presenting the steps of a procedure in a clear and organized fashion. It is interesting to note that only 56% of lessons contained segments that scored above Low on *Connections Across Representations*. While this seems like comparatively few lessons (particularly compared to the percentage of lessons that scored above Low on most other dimensions), it is important to note that some content within the algebra curriculum may lend itself more readily to this practice. For example, a lesson on multiplying binomials may by design include only the symbolic representation and therefore not provide opportunities for connecting to other representational forms. In other words, there may be good reasons why teachers do not engage in this (or other) instructional practices. In fact, while this practice is one that has been shown to benefit student understanding (Knuth, 2000), there may be situations where engaging in this particular practice might be cumbersome or inappropriate.

Notably, in the lessons where this instructional feature was present, it appears to have been instantiated only briefly or without much depth.

More striking, however, is that only 39% of lessons included a segment that scored above Low on *Situating the Mathematics*. This indicates that over two thirds of lessons contained no instances of teachers making connections to other aspects of the algebra curriculum, to the broader domain of mathematics, or motivating the current area under study within a larger mathematical context. This code should not be content-dependent in the same way as *Connecting Across Representations* might be. That is, regardless of content, it is still viable that teachers might connect the topic under study in the lesson to other aspects of the algebra curriculum or to other mathematical ideas outside of algebra. While this study does not give insight into why teachers engage in this practice infrequently, supporting teachers in developing and deepening this practice may help students to connect prior knowledge with new content and adapt and expand their prior understandings (Kieran, 2013).

Looking more closely at typical lessons in the sample, I found that these lessons featured glimmers of promising practices that could be capitalized upon and deepened, perhaps to the benefit of students. That these practices appear occasionally and briefly implies that teachers may be aware of these practices and are beginning to engage in them. With attention and development, it is reasonable to believe that they can be elevated to higher levels of quality. For example, in lessons in which teachers support procedural flexibility by discussing multiple pathways through a mathematical procedure briefly, but without elaboration, it is possible to envision building capacity so that instruction might be deepened to attend more explicitly to supporting procedural

flexibility. I also found that many lessons contained missed opportunities to engage in various practices identified in the QIPA. This was not an assessment that teachers *should* be engaging in particular practices but were not, but rather an observation that teachers frequently opened the door through their instruction to enable the enactment of particular instructional practices, but did not follow through in engaging in them. Making these missed opportunities explicit could allow teachers to capitalize on what they are already doing to deepen their practice.

The findings from this study should not be used to make general and deterministic claims about algebra instruction or algebra teachers, as teachers who participated in the MET study are a non-representative sample of ninth grade algebra teachers. Districts, schools, and teachers volunteered to participate in the study and teachers selected their own lessons for videotaping. Thus, there may be systematic differences between teachers who chose to participate and those who did not, and the observed instruction may not be representative of all teachers. Within teachers, video-recorded lessons may not be representative of their regular practice. In addition, I do not claim that the instructional practices highlighted in this paper are the best or the only practices worthy of focus and attention. This study's design, for example, does not allow me to make claims that deepening these particular practices will improve student achievement. Yet by better understanding the nature and quality of current instructional practice, we open up the possibilities for a discussion on improving instruction.

The lessons learned from a close examination of instructional practice from a sample this large should thus help not only inform future research on more representative samples, but reveal potential avenues for instructional improvement. The results of this

study can be used to frame a discussion on how we as a field can work with teachers to develop and deepen particular discipline-specific instructional practices that show promise for student learning.

Taken together, the findings presented in this paper together indicate that while instruction does not feature the practices identified in the QIPA in depth or at high levels of quality, it is important to note that they *do* seem to occur at more superficial levels across many lessons in this sample. Thus, the description presented above can be interpreted in two ways—the first is one of despair in which high quality instructional practices are largely absent in algebra lessons in this sample. This is not necessarily a useful interpretation if the goal is instructional improvement. The other interpretation is one of cautious optimism. It may be in fact that teachers are engaging in these practices at low levels and that this provides opportunities to deepen their enactment.

As states and districts begin to implement the Common Core State Standards and prepare for new, demanding assessments, there is a push for teachers to focus their instruction in ways that allow students to become proficient with and make meaning of the mathematics they encounter. I argue that in order to improve instruction we must not only understand what instructional practices we hope to see, but also what instructional practices we most often *do* see, and at what level of quality they are enacted. In doing so, we not only get a better sense of where instruction is relative to where we might wish it to be, but we are better able to build upon existing skills and practices, engaging teachers in the work of instructional improvement.

Appendix A
Distribution of Sample Teachers Across MET Districts

Table A1

Distribution of sample teachers across MET districts

| District | Full MET sample | Full Sample of 9 th Grade Math teachers (n=233) | 9 th Grade Teachers with CKT Scores (n=141) | 9 th Grade Math teachers with CKT and video (n=81) |
|----------|-----------------|--|--|---|
| A | 26.1 | 18.9 | 17.8 | 19.8 |
| B | 4.9 | 13.3 | 14.9 | 14.8 |
| C | 15.5 | 16.3 | 14.2 | 17.3 |
| D | 16.3 | 12.0 | 12.8 | 17.3 |
| E | 26.1 | 39.5 | 40.4 | 30.9 |

Note. 11.0% of teachers in full sample are from District E, which did not include ninth grade algebra teachers in the sample.

Appendix B

Quality of Instructional Practice in Algebra (QIPA) Scoring Guide

TEACHING PROCEDURES

*This domain is intended to capture the quality of instruction on procedures. We define procedures here as instructions for completing a mathematical algorithm or task. These codes refer to **the presentation of new procedures, review of previously learned material, and descriptions of procedures used in the context of solving a problem/problems**. This can occur during teacher direct instruction or while the teacher is interacting with students during independent or group work time. Do not count students quietly practicing procedures here.*

Note that calculator instructions do not count as a procedure, but using the calculator in service of a mathematical procedure is acceptable and should be coded in the context of the mathematical procedure being undertaken. Instruction that focuses solely on how to manipulate the calculator does not by itself count as a procedure.

This domain consists of the following codes:

Making Sense of Procedures
Supporting Procedural Flexibility
Organization in the Presentation of Procedures
Overall Teaching of Procedures

Before scoring the individual codes, for all segments, please indicate whether any part of the segment focused on the teaching of procedures as defined above.

Is a Procedure Taught in this Segment:

| NO | YES |
|--|---|
| The segment does not include instruction on procedures (e.g. may be entirely quiet student work time; instruction on a concept; inquiry or exploration that does not include procedures, etc.) | The segment includes instruction on procedures. |

MAKING SENSE OF PROCEDURES

In making sense of a procedure, a teacher or students may do one or more of the following (not exhaustive):

- *Make meaning of individual steps in the procedure (e.g. why you plug in $x=0$ into a linear equation when finding the y-intercept)*
- *Make meaning of the solution generated by the procedure (e.g. the values of x and y in the solution to a system of linear equations are a coordinate pair that tell you the point of intersection of the two lines)*
- *Attend to the purpose/mathematical goal of the procedure (e.g. using quadratic formula allows us to find the roots of a parabola)*
- *Attend to the mathematical properties underlying a procedure (e.g. How FOIL is really the distributive property)*
- *Attend to why a procedure holds (e.g. when you multiply exponents with a common base, you can add the exponents because multiplication works as repeated addition)*
- *Make sense of the overall procedure (e.g. teacher develops students' intuition about how to graph linear inequalities by making sense of the solution set as a region of points that makes the inequality true)*

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|--|---|---|
| <p>Not present—no instruction on procedures occurs.</p> <p>OR</p> <p>Procedure is presented with no attention to meaning or sense-making (e.g. as a recipe to be followed with a narrow application (when equations look like this... do that...)).</p> <p>OR</p> <p>Errors in the presentation of the procedure muddle the opportunity for students to make sense of the procedure.</p> | | <p>Teacher or students make sense of the procedure(s) more than briefly (e.g. several instances within a segment or one somewhat more sustained instance), but it is not the focus of the instruction on procedures.</p> <p>Also score here if the segment includes features of High but also some small errors or imprecisions that impact meaning or sense-making.</p> | | <p>Teacher makes sense of the procedure(s) throughout the instruction by:</p> <p>a) a combination of elements that saturate the segment with meaning or sense-making</p> <p>b) a sustained focus on one of the elements of sense-making</p> |

SUPPORTING PROCEDURAL FLEXIBILITY

This code captures the degree to which teachers present algebraic procedures to students in a way that affords students the opportunity to begin to develop procedural flexibility, for example identifying what procedures to apply and knowing when to apply them.

In supporting flexibility, the teacher or students may:

- *Note multiple pathways through a procedure*
- *Attend to applicability conditions of a procedure (e.g. by noting when it can be used or what led to the choice of a given procedure)*
- *Attend to key conditions of steps within a procedure (e.g. by noting what must hold for a procedure to work or for a particular step of the procedure to be executed, pointing to key decision points).*
- *Use a procedure to conjecture about general rules*
- *Compare multiple procedures for their affordances/limitations*

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|---|---|--|
| <p>Not present—no instruction on procedures occurs.</p> <p><i>OR</i></p> <p>Procedure is presented without any attention to elements of flexibility.</p> <p><i>OR</i></p> <p>Presentation of the procedure is error-filled to the extent that it hinders the development of flexibility.</p> | | <p>Teacher or students attend to flexibility more than briefly, but do not elaborate or it is not the focus of the instruction on procedures.</p> | | <p>Teacher or students explicitly attend to flexibility in a sustained way throughout the segment by:</p> <p>a) Attention with some elaboration to multiple elements of flexibility within the segment</p> <p><i>OR</i></p> <p>b) A sustained focus on one aspect of flexibility that characterizes the segment.</p> |

ORGANIZATION IN THE PRESENTATION OF PROCEDURES

This code indicates how complete, detailed, correct, and organized the teacher's (or students') presentation of content is when outlining or describing procedures, or describing the steps of a procedure used to solve problems.

Note: This code does not require any meaning to be made of the procedure(s). Being complete, organized and clear can happen regardless of the meaning-making orientation of the instruction. Rather this code focuses on organization and clarity of the presentation or procedures—either verbally or on the board (or in combination)—as well as how teachers (or students) highlight and clarify necessary information.

Note: Score down for errors that hinder mathematical clarity of the presentation of the procedure (e.g. overdetermination of the conditions for use of a procedure such as stating that in order to graph a line, the equation must be in standard form).

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|--|---|--|
| <p>There are no examples of presentation of procedures.</p> <p>OR</p> <p>The teacher's presentation of the procedure is disorganized, incorrect, incomplete, or unclear.</p> | | <p>The teaching of the procedure is acceptable, complete, and mostly clear, but not exceptionally organized or detailed.</p> | | <p>The teaching of the procedure is not only clear, complete, and free of errors, but is exceptionally organized and/or detailed.</p> <p>This may include:</p> <ul style="list-style-type: none"> • Exceptionally careful and systematic organization of the material either verbally or in writing • Explicit reminders of steps in the procedure to pay attention to (e.g. common mistakes at key junctures) • Generalization of procedures beyond specific problems. <p>These features occur in a sustained way and/or characterize the segment.</p> |

OVERALL TEACHING OF PROCEDURES

This code captures the overall quality of the instruction on procedures.

Note: This code is a holistic code for each segment. It is not an average of the 3 codes in this domain, rather it is an overall estimate of the quality of the instruction on procedures.

In scoring Overall Teaching Procedures, we reserve a score of 1 for instances where Teaching Procedures did not occur or all components were either unclear or incorrect. If there are scores above 1 on any of the individual codes in this domain, the Overall Teaching Procedures code cannot be a 1.

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|---|---|--|
| <p>There is no instruction on procedures</p> <p>OR</p> <p>Procedures are taught but incorrect or full of errors.</p> | | <p>The teaching of procedures is of reasonable quality as indicated by some combination of meaning-making or attention to flexibility and/or reasonable organization.</p> | | <p>The teaching of procedures is characterized by one of the following:</p> <ul style="list-style-type: none"> a) Outstanding performance in one or more of the codes of this domain b) A combination of strong elements across the codes in this domain that together contribute to the quality of the teaching of procedures in the segment. |

LEVERAGING CONNECTIONS

This domain is intended to capture the connections that teachers and students forge between and among algebraic concepts, representations, and abstract ideas.

This domain consists of the following codes:

Connecting Across Representations

Situating the Mathematics

Making Connections Between Concrete and Abstract Ideas in Algebra

Overall Leveraging Connections

Note: The practices in this domain occur in multiple contexts and thus these codes are scored for all segments, regardless of whether instruction on procedures is occurring in the segment.

CONNECTING ACROSS REPRESENTATIONS

*This code captures the nature of the connections teachers or students make between and across representational forms of **graph**, **table**, **equation/symbolic**, and **context** of algebraic problems, ideas, and concepts.*

For this code to be scored a 4 or 5, more than one representation must be visually or verbally present (e.g. a word problem read aloud). A segment in which a teacher references a graphical representation while discussing an equation, but in which the graph is not visible or the teacher does not explicitly describe it cannot score above 3. If you have seen a graph in a prior segment but do not see it in the current segment, you may count it as visible. A teacher using an algebraic procedure (manipulating an equation) to solve a word problem (context) would score 4 or higher only if the problem context had been explicitly stated verbally or in writing and if explicit connections were made between elements of the algebra and aspects of the context.

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|--|---|---|
| <p>Not present—only one representation is present or more than one representation is used but no connections are made.</p> <p>OR</p> <p>Teacher (or students) makes connections across representations that are incorrect.</p> | | <p>Teacher (or students) more than briefly reference a connection between or across representations but with no further elaboration. (e.g. using a graph to locate the y-intercept of a line and using that value to write the linear equation without additional discussion).</p> | | <p>Teacher (or students) make explicit, careful connections between and across representations (e.g. showing how a change in one representation manifests in another)</p> |

SITUATING THE MATHEMATICS

Teacher or students make connections to other aspects of the algebra curriculum, related topics, or the broader domain of mathematics, situating and motivating the current area under study within a broader context. Teachers/students may:

- *Make connections to prior content by linking what the class is currently doing to other topics the class has studied (e.g. Now that we have eliminated one variable, this equation looks like the single-variable algebra equations we already know how to solve)*
- *Make connections to future content by articulating how what the class is currently learning sets up other topics under study (e.g. Knowing how graph a line is necessary to be able to later graph inequalities).*
- *Use the architecture of the lesson to develop a mathematical through line. (e.g. A teacher articulates the mathematical role of the Do Now in preparing for or motivating the main topic of the lesson or explicitly uses the solution of the Do Now in the context of the lesson. Note here we do not intend to capture instances where the teacher simply references the Do Now in the lesson, but rather when there is explicit building of mathematical knowledge from the Do Now.)*
- *Make connections to other mathematical topics or the broader domain of mathematics (e.g. noting how changing the coefficients in a particular type of function work similarly across families of functions; relating finding area of geometric shapes to area under a curve).*

| 1 | 2 | 3 (modal) | 4 | 5 |
|---|---|---|---|---|
| <p>Not present—no explicit situating of the mathematics happens in the segment.</p> <p>OR</p> <p>Teacher situates the mathematics in a way that is incorrect.</p> | | <p>Teacher (or students) references a connection that situates the mathematics more than briefly but with little elaboration and without the features under High.</p> | | <p>Teacher (or students) make connections that situate the mathematics, drawing correspondences and/or motivating the topic under study, characterized by at least one of the following:</p> <p>a) Explicitness about how topics or ideas are connected</p> <p>b) Elaboration and detail about the nature of the connection</p> <p>These features occur in a sustained way and/or characterize the segment.</p> |

CONNECTING BETWEEN CONCRETE & ABSTRACT IDEAS

This code is intended to capture the degree to which the teacher or students leverage concrete examples, representations or ideas to develop understanding of abstract concepts, formulas, notation, and definitions. We define abstraction here as mathematical ideas expressed in theoretical or symbolic form, formulas and algebraic rules, and algebraic concepts, definitions, or ideas.

They may do one or more of the following (this list is not exhaustive):

- *Unpack the components of a formula by attending to what the symbols in the formula represent (e.g. identifying the initial value and common difference in an arithmetic sequence or discussing what the components of the slope-intercept formula represent)*
- *Represent an abstract situation or idea symbolically (e.g. making sense of what variables mean in a contextualized situation)*
- *Clarify mathematical definitions or abstract algebraic concepts using concrete examples (e.g. having students draw examples and counter-examples of supplementary angles)*
- *Leverage abstract algebraic tools or representations to solve concrete examples (e.g. to find the distance between two buildings on a city block, one could superimpose a right triangle and use the Pythagorean Theorem to find the distance).*
- *Explicitly relate an abstract concept to its analogous concrete idea (e.g. demonstrating the relationship between factoring a polynomial and factoring a numerical product)*
- *Use concrete, or pictorial examples or manipulatives to introduce or illustrate abstract ideas or concepts (e.g. algebra tiles to model combining like terms)*
- *Use numbers to develop generalized rules and properties (e.g. develop the rule for negative exponents by exploring 3^4 , 3^3 , 3^2 , 3^1 , 3^0 , 3^{-1} , 3^{-2} , etc.)*

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|---|---|---|
| <p>Not present—abstract concepts definitions, formulas or notation are not taught/no connections are made.</p> <p>OR</p> <p>Teacher (or students) incorrectly makes connections between concrete and abstract examples, representations and ideas.</p> | | <p>Teacher (or students) make more than brief connections but these connections are not fully elaborated and are not a substantial focus of the segment.</p> <p>E.g., a teacher may explicitly label and briefly comment on what the numbers in a formula represent but this is done without further elaboration.</p> | | <p>Teacher (or students) make these connections in a sustained way throughout the segment. Connections are explicit and elaborated upon and:</p> <p>a) Are the focus of the segment, occurring in a sustained way throughout</p> <p>OR</p> <p>b) Occur for only a portion of the segment but is the major feature of the teacher-student work</p> |

OVERALL LEVERAGING CONNECTIONS

This code captures the overall quality of the connections forged in the segment.

Note: This code is a holistic code for each segment. It is not an average of the 3 codes in this domain, rather it is an overall estimate of the quality of the connections made in the segment.

In scoring this code, we reserve a score of 1 for instances where the codes in this domain were not present or all were either unclear or incorrect. If there are scores above 1 on any of the individual codes in this domain, the Overall Leveraging Connections code cannot be a 1.

| 1 | 2 | 3 (modal) | 4 | 5 |
|--|---|---|---|--|
| <p>Not Present— There are no connections across representations, no instances of situating the mathematics and no connections between concrete and abstract ideas, representations or examples</p> <p>OR</p> <p>Connections happen but are incorrect or unclear.</p> | | <p>Some types of connections are more than minimally present but these do not characterize the segment.</p> <p>For example, a segment may contain one or more Mid scores in the codes of this domain; a number of brief instances that together add up to a Mid; or a mix of stronger and weaker connections.</p> | | <p>Connections are present and meaningful in the segment as characterized by:</p> <p>a) Outstanding performance in one or more of the codes of this domain</p> <p>b) A combination of strong elements across the codes in this domain that together build meaning from the connections forged.</p> |

WHOLE-LESSON FORMAT OF INSTRUCTION CODES

INQUIRY/EXPLORATION

*This whole-lesson code is intended to capture the amount of exploration and inquiry **in the lesson as a whole**. We define inquiry/exploration as tasks in which students are asked to do significant mathematical work, often without a pre-determined solution or solution path. This type of task generally features students working in groups or teams to complete an extended task with little teacher-directed instruction and/or scaffolding.*

| 1 (not true of this lesson) | | 3 (somewhat true of this lesson) | 4 | 5 (very true of this lesson) |
|--|--|---|---|---|
| Not present—students spend no time in the lesson on exploration or inquiry into algebraic concepts | | The lesson features <i>some</i> student exploration and/or some work on a somewhat cognitively demanding task | | The majority of the lesson is taken up by student exploration or inquiry into an algebraic concept and/or the student exploration or inquiry characterizes the lesson |

TEACHER-LED INSTRUCTION

This whole-lesson code is intended to capture the degree to which the teacher directs the content and processes in the lesson and the degree to which students actively engage in doing mathematics in the lesson.

| 1 (not true of this lesson) | 2 | 3 (somewhat true of this lesson) | 4 | 5 (very true of this lesson) |
|---|---|--|---|--|
| Not present—instruction is mostly student-centered with minimal direction from the teacher. | | Instruction is largely teacher directed though may contain some time for independent/group student practice or inquiry. Students participate in the lesson to some degree with meaningful mathematical contributions and/or time to work through concepts on their own, but instruction is largely teacher-directed. | | Teacher directs all content for the duration of the lesson with no student inquiry and little-to-no student practice |

Appendix C Case Study Protocol

Adapted from Yin, 2009

Guiding Questions:

- *How is [specific practice] instantiated in this lesson?*
- *How can I describe this practice at varying levels of quality in this lesson?*
- *How is this lesson illustrative of lessons in the sample?*

Data Collection:

Teacher ID: _____ Lesson ID: _____

Lesson Observation Notes:

| | | |
|--|--------------------------------------|--|
| Description of notable/typical instance of [specific practice] | <i>What teacher is doing/saying:</i> | <i>What students are doing/saying:</i> |
| | | |
| Time Stamp: | | |
| Description of notable/typical instance of [specific practice] | <i>What teacher is doing/saying:</i> | <i>What students are doing/saying:</i> |
| | | |
| Time Stamp: | | |
| Description of notable/typical instance of [specific practice] | <i>What teacher is doing/saying:</i> | <i>What students are doing/saying:</i> |
| | | |
| Time Stamp: | | |

Case Study Questions [sources of data]:

1. Describe [specific practice] in detail in this lesson, including specific instances and notable features. [Video observation]
2. What aspects of [specific practice] in this lesson context contribute to the lesson's coded score on this dimension. [video observation, scores on [specific practice] code]
3. In what ways is this lesson "typical" of [specific practice] in this sample? Describe any illustrative vignettes in this lesson. [video observation, descriptive quantitative results]
4. In what ways is this lesson atypical of [specific practice] in this sample? Describe any illustrative vignettes in this lesson. [video observation, descriptive quantitative results]
5. In what ways does the content of the lesson lend itself to [specific practice]?
6. What other instructional practices seem to interact with [specific practice] in this lesson? In what ways? [video observation, scores on [specific practice] code, scores on other codes from other instruments]
7. How do teacher and students engage around [specific practice] in this lesson? [video observation]
8. Are there missed opportunities to engage in [specific practice]?
9. How do instantiations of [specific practice] in this lesson interact with other sources of data about the lesson and classroom? [video observation]
10. Are there patterns emerging across lessons? [video observation, other case study protocols, analytic memos]

References

- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: United States Department of Education.
- Allensworth, E., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College preparatory curriculum for all: Academic consequences of requiring Algebra and English I for ninth graders in Chicago. *Educational Evaluation and Policy Analysis*, 31(4), 367–391.
- Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Chapel Hill, NC: Horizon Research, Inc.
- Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, 17(2–3), 62–87.
- Bill & Melinda Gates Foundation. (2010a). *Classroom observations and the MET project*. Author. Retrieved from www.metproject.org
- Bill & Melinda Gates Foundation. (2010b). *Content knowledge for teaching and the MET project*. Author. Retrieved from www.metproject.org
- Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *The Teachers College Record*, 110(3), 608–645.
- Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction*, 25, 24–34.
- Booth, L. R. (1988). Children's difficulties in beginning algebra. In A. F. Coxford, & A. P. Shulte (Eds.), *The ideas of algebra, K–12*. 1988 Yearbook of the National Council of Teachers of Mathematics (pp. 299–306). Reston, VA: National Council of Teachers of Mathematics.
- Borasi, R., & Fonzi, J. (2002). *Professional development that supports school mathematics reform* (Foundations Monograph, Vol. 3). Arlington, VA: National Science Foundation.

- Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Whittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.
- Brown, L. (2004). Responses to ‘the core of algebra:’ The challenge of learning through transformational algebraic activity. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The future of teaching and learning of algebra: The 12th ICMI study* (pp. 35–40). Norwell, MA: Kluwer Academic Publishers.
- Brown, J., Schiller, K., Roey, S., Perkins, R., Schmidt, W., & Houang, R. (2013). *Algebra and geometry curricula* (NCES 2013-451). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19–32). New Jersey: Lawrence Erlbaum Associates.
- Charalambous, C. Y., & Hill, H. C. (2012). Teacher knowledge, curriculum materials, and quality of instruction: Unpacking a complex relationship. *Journal of Curriculum Studies*, 44(4), 443–466.
- Chazan, D. (2000). *Beyond formulas in mathematics and teaching: Dynamics of the high school algebra classroom*. New York: Teachers College Press.
- Chazan, D. (2008). The shifting landscape of school algebra in the United States. In C. Greenes & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: Seventieth yearbook* (pp. 19–33). Reston, VA: National Council of Teachers of Mathematics.
- Chazan, D., & Yerulshlami, M. (2003). On appreciating the cognitive complexity of school algebra: Research on algebra learning and directions for curricular change. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A Research companion to principles and standards for school mathematics* (pp. 123–135). Reston, VA: National Council for Teachers of Mathematics.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2012). *The aftermath of accelerating algebra: Evidence from a district policy initiative* (National Bureau of Economic Research, Working Paper 18161). Retrieved from: <http://www.nber.org/papers/w18161>
- Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, 12(3), 311–329.

- Cohen, D. K. (2011). *Teaching and its predicaments*. Cambridge, MA: Harvard University Press.
- Cohen, D. K., & Hill, H. C. (2001). *Learning policy*. New Haven, CT: Yale University Press.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25(2), 119–142.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory, third edition*. Thousand Oaks, CA: Sage Publications.
- Cortes, K., Goodman, J. & Nomi, T. (2015). Intensive math instruction and educational attainment: Long-run impacts of double-dose algebra. *Journal of Human Resources*, 50, 108–158.
- Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms, 1890–1990*. New York: Teachers College Press.
- Cuban, L. (2013). *Inside the black box of classroom practice*. Cambridge, MA: Harvard Education Press.
- Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *The Journal of Mathematical Behavior*, 15(4), 375–402.
- Desimone, L. (2011). R&D: A primer on effective professional development. *Phi Delta Kappan*, 92(6), 68–71.
- Diamond, J. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education*, 80(4), 285–313.
- Doerr, H. (2004). Teachers knowledge and the teaching of algebra. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The future of the teaching and learning of algebra* (pp. 267–290). Norwell, MA: Kluwer Academic Publishers Group.
- Gamoran, A., & Hannigan, E. (2000). Algebra for all? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis*, 22(3), 241–254.
- Gardner, D. (1983). *A nation at risk: The imperative for educational reform—Report of the National Commission on Excellence in Education*. Washington, DC: U.S. Government Printing Office.

- Gargani, J., & Strong, M. (2014). Can we identify a successful teacher better, faster, and cheaper? Evidence for innovating teacher observation systems. *Journal of Teacher Education*, 65(5), 389–401.
- Gitomer, D. H., Bell, C. A., Qi, Y., McCaffrey, D. F., Hamre, B. K., & Pianta, R. C. (2014). The instructional challenge in improving teaching quality: Lessons from a classroom observation protocol. *Teachers College Record*, 116(6), 1–32.
- Gitomer, D. H., Phelps, G., Weren, B., Howell, H., & Croft, A. J. (2014). Evidence on the validity of content knowledge for teaching assessments. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 493–528). San Francisco, CA: Jossey-Bass.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Graham, P. (2005). *Schooling in America*. New York: Oxford University Press.
- Grossman, P. L., Loeb, S., Cohen, J., & Wykoff, J. (2013). Measure for measure: The relationship between measures of instruction practice in middle school English Language Arts and teachers' value-added scores. *American Journal of Education*, 119(3), 445–470.
- Grossman, P. L., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher*, 24(5), 5–11, 23.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82.
- Guiron, G., & Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis*, 17(3), 323–336.
- Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice*. New York: Taylor & Francis Group, LLC.
- Hacker, A., (2012, July 28). Is algebra necessary? *The New York Times*. Retrieved from <http://www.nytimes.com>
- Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 U.S. early childhood and elementary classrooms*. New York: Foundation for Child Development.
- Hiebert, J. (1999). Relationships between research and the NCTM standards. *Journal for Research in Mathematics Education*, 30(1), 3–19.

- Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 5–24). Reston, VA: National Council for Teachers of Mathematics.
- Hiebert, J. (2013). The constantly underestimated challenge of improving mathematics instruction. In K. R. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 45–56). New York: Springer.
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Reston, VA: National Council of Teachers of Mathematics.
- Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M., ..., & Gallimore, R. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis*, 27(2), 111–132.
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., ..., & Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment*, 17(2–3), 88–106.
- Hill, H. C., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, 83(2), 371–384.
- Hill, H. C., Litke, E., Lynch, K., Pollard, C., & Gilbert, B. (2014, September). Learning lessons from instruction: Descriptive results from an observational study of urban elementary classrooms. Paper presented at Association for Public Policy and Management, Segovia, Spain.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Ho, A. D., & Kane, T. J. (2013). *The reliability of classroom observations by school personnel* (Research Paper). Seattle, WA: Measures of Effective Teaching Project, Bill & Melinda Gates Foundation.

- Jacobs, J. K., Hiebert, J., Givvin, K. B., Hollingsworth, H., Garnier, H., & Wearne, D. (2006). Does eighth-grade mathematics teaching in the United States align with the NCTM standards? Results from the TIMSS 1995 and 1999 video studies. *Journal for Research in Mathematics Education*, 37(1), 5–32.
- Jacobs, J. K., Kawanaka, T., & Stigler, J. W. (1999). Integrating qualitative and quantitative approaches to the analysis of video data on classroom teaching. *International Journal of Educational Research*, 31(8), 717–724.
- Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, 23(1), 51–68.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: American Council on Education/Praeger.
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* (Research Paper). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from <http://www.metproject.org>
- Kaput, J. J. (1995). Long term algebra reform: Democratizing access to big ideas. In C. Lacampagne, J. Kaput, & W. Blair (Eds.), *The Algebra initiative colloquium*, (Vol. 1, pp. 33–49). Washington, DC: Department of Education, Office of Research.
- Kaput, J. J. (1999). Teaching and learning a new algebra. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 133–155). New Jersey: Lawrence Erlbaum Associates.
- Kieran, C. (1992). The learning and teaching of school algebra. In D. Grouws. (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 390–419). New York: Macmillan Publishing Co., Inc.
- Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels. In F. Lester (Ed.), *The second handbook of research on mathematics teaching and learning* (pp. 707–762). Reston, VA: National Council for Teachers of Mathematics.
- Kieran, C. (2013). The false dichotomy in mathematics education between conceptual understanding and procedural skills: An example from algebra. In K. R. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 153–171). New York: Springer.

- Kilpatrick, J. & Izak, A. (2008). A history of algebra in the school curriculum. In C. Greenes, & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: Seventieth yearbook* (pp. 3–18). Reston, VA: National Council of Teachers of Mathematics.
- Knuth, E. J. (2000). Student understanding of the Cartesian connection: An exploratory study. *Journal for Research in Mathematics Education*, 31(4), 500–507.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37–62.
- Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, 14(1), 25–47.
- Litke, E. (2014). Beyond measuring effectiveness: A descriptive look at algebra instruction in six urban districts. (Unpublished qualifying paper). Harvard Graduate School of Education, Cambridge, MA.
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Lynch, K., & Star, J. R. (2014a). Teachers' views about multiple strategies in middle and high school mathematics. *Mathematical Thinking and Learning*, 16(2), 85–108.
- Lynch, K., & Star, J. R. (2014b). Views of struggling students on instruction incorporating multiple strategies in Algebra I: An exploratory study. *Journal for Research in Mathematics Education*, 45(1), 6–18.
- Marder, M., Walkington, C., Abraham, L., Allen, K., Arora, P., Daniels, M., & Walker, M. (2010). *The UTeach Observation Protocol (UTOP) training guide (adapted for video observation ratings)*. Austin, TX: UTeach Natural Sciences, University of Texas Austin.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., ... & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732–749.
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*, 3rd Edition. Thousand Oaks, CA: Sage Publications.
- McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32(5), 493–517.

- McCrorry, R., Floden, R., Ferrini-Mundy, J., Reckase, M., & Senk, S. L. (2012). Knowledge of algebra for teaching: A framework for knowledge and practices. *Journal for Research in Mathematics Education*, 43(5), 584–615.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco: Jossey-Bass.
- Moses, R. P. & Cobb, C. E. J. (2001). *Radical equations: Organizing math literacy in America's schools*. Boston, MA: Beacon Press.
- National Council for Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council for Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Council for Teachers of Mathematics. (2008). *Algebra: What, when and for whom*. Reston, VA: Author. Retrieved from www.nctm.org
- National Council for Teachers of Mathematics. (2013). *Supporting the Common Core State Standards for mathematics*. Reston, VA: Author. Retrieved from www.nctm.org
- National Council for Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards (mathematics)*. Washington, DC: Author.
- National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.
- Oakes, J. (1990). *Multiplying inequalities: The effects of race, class, and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: Rand Corporation.
- O'Connor, M. C. (1998). Language socialization in the mathematics classroom: Discourse practices and mathematical thinking. In M. Lampert & M. Blunk (Eds.), *Talking mathematics: Studies of teaching and learning in school* (pp. 17–55). New York: Cambridge University Press.
- Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007) *Organizing instruction and study to improve student learning* (NCER

- 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ncer.ed.gov>
- Payne, C. M. (2008). *So much reform, so little change: The persistence of failure in urban schools*. Cambridge, MA: Harvard Education Press.
- Pianta, R., & Hamre, B. (2009). Conceptualization, measurement, and improvement of classroom processes: Standard observation can leverage capacity. *Educational Researcher*, 38(2), 109–119.
- Piburn, M., Sawada, D., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed teaching observation protocol (RTOP)*. Arizona Collaborative for Excellence in the Preparation of Teachers.
- Porter, A. C., Floden, R., & Fuhrman, S. (1998). The effects of upgrading policies on high school mathematics and science. *Brookings Papers on Education Policy*, No. 1, 123–172.
- Porter, A. C., & Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (100th Yearbook of the National Society for the Study of Education, part II, pp. 60–80). Chicago: National Society for the Study of Education.
- Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school*. Boston: Houghton Mifflin Company.
- President's Council of Advisors on Science and Technology. (US). (2010). *Prepare and inspire: K-12 education in science, technology, engineering, and math (STEM) for America's future: Executive report*. Washington, DC: Executive Office of the President, President's Council of Advisors on Science and Technology.
- Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research*, 80(3), 372–400.
- Raymond, A., & Leinenbach, M. (2000). Collaborative action research on the learning and teaching of algebra: A story of one mathematics teacher's development. *Educational Studies in Mathematics*, 41(3), 283–307.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93(2), 346–362.

- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, 101(3), 529–544.
- Schoenfeld, A. (1995). Report of working group 1. In C. Lacampagne, W. Blair, & J. Kaput (Eds.), *The Algebra initiative colloquium: Papers presented at a conference on reform in algebra*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Skemp, R. R. (2006). Relational understanding and instrumental understanding. *Mathematics teaching in the middle school*, 12(2), 88–95.
- Smith, J.B., Lee, V.E., & Newmann, F.M. (2001). *Improving Chicago's schools: Instruction and achievement in Chicago elementary schools*. Chicago: Consortium on Chicago School Research.
- Smith, T. M. (2014). Instructional practices to support student success in Algebra I. U.S. Department of Education. Retrieved from: <http://www2.ed.gov/programs/dropout/instructionalpractices092414.pdf>
- Spillane, J. P., & Zeuli, J. S. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis*, 21(1), 1–27.
- Stacey, K., & Chick, H. (2004). Solving the problem with algebra. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The future of the teaching and learning of algebra* (pp. 1–20). Norwell, MA: Kluwer Academic Publishers Group.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, 36(5), 404–411.
- Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., ..., & Zbiek, R. M. (2015). *Teaching strategies for improving algebra knowledge in middle and high school students* (NCEE 2014-4333). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from: <http://whatworks.ed.gov>
- Star, J. R., Herbel-Eisenmann, B. A., & Smith III, J. P. (2000). Algebraic concepts: What's really new in new curricula? *Mathematics Teaching in the Middle School*, 5(7), 446–451.

- Star, J. R., & Rittle-Johnson, B. (2009). Making algebra work: Instructional strategies that deepen student understanding, within and between algebraic representations. *ERS Spectrum*, 27(2), 11–18.
- Straus, V. (2014, November 8). Why so many parents are freaking out about Common Core math. *The Washington Post*. Retrieved from <http://www.washingtonpost.com>
- Stein, M.K., Kaufman, J., Sherman, M., & Hillen, A. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of Educational Research*, 81(4), 453–492.
- Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, 2(1), 50–80.
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. Lester (Ed.), *The second handbook of research on mathematics teaching and learning* (pp. 319–370). Reston, VA: National Council for Teachers of Mathematics.
- Stein, M. K., Smith, M., Henningsen, M. & Silver, E. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development*. New York: Teachers College Press.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. New York: The Free Press.
- Stodolsky, S. S. (1988). *The subject matters*. Chicago: University of Chicago Press.
- Tate, W. F. (1997). Race-ethnicity, SES, gender and language proficiency trends in mathematics: An update. *Journal for Research in Mathematics Education*, 28(6), 652–679.
- Thompson, P. W. (2013). In the absence of meaning... In K. R. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 57–93). New York: Springer.
- Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. In A. F. Coxford, & A. P. Shulte (Eds.), *The ideas of algebra, K-12*. 1988 Yearbook of the National Council of Teachers of Mathematics (pp. 7–13). Reston, VA: National Council of Teachers of Mathematics.
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89–122.

Yin, R. K. (2009). *Case study research: Design and methods, fourth edition*. Thousand Oaks, CA: Sage Publications.

Tables

Table 1

Baseline Characteristics of Sample Teachers

| | All 9 th grade teachers Y1 (n=233) | 9 th Grade Teachers with CKT Scores (n=141) | 9 th Grade Math teachers with CKT and video (n=81) |
|---|--|--|---|
| Percent Female | 56.7 (8% missing) | 57.5 (7% missing) | 60.5 (5% missing) |
| Percent White | 52.8 | 55.3 | 65.4 |
| Percent Black | 27.9 | 27.7 | 22.2 |
| Percent Hispanic | 5.2 | 5.7 | 6.2 |
| Percent Other Race | 6.0 | 4.3 | 1.2 |
| Mean Years of Experience (Range)* | 7.5 (0–35) <i>n</i> =75 | 6.5 (0–33) <i>n</i> =46 | 6 (0–24) <i>n</i> =28 |
| Mean Years of Experience in District (Range)* | 6.1 (0–35) <i>n</i> =185 | 5.9 (0–33) <i>n</i> =113 | 5.7 (0–24) <i>n</i> =63 |
| Percent Masters Degree or More* | 27.9 <i>n</i> = 139 | 26.3 <i>n</i> =83 | 32.1 <i>n</i> =56 |
| CKT | | | |
| Mean (sd) | | 61.7 (14.3) | 62.7 (14.22) |
| Median | | 62.9 | 62.9 |
| Min score | | 22.9 | 22.9 |
| Max score | | 97.1 | 97.1 |
| Range of scores | | 74.2 | 74.2 |

*Reported for teachers for whom this information was available

Table 2

*Descriptions of the Codes in the Teaching Procedures Domain of QIPA***Teaching Procedures**

These codes are intended to capture the quality of instruction on procedures

| | |
|---|--|
| Making Sense of Procedures | <p>The degree to which a teacher (or students) engage in sense-making in the context of teaching procedures by one or more of the following:</p> <ul style="list-style-type: none"> • Make meaning of individual steps in the procedure • Make meaning of the solution generated by the procedure • Attend to the purpose/mathematical goal of the procedure • Attend to the mathematical properties underlying a procedure • Attend to why a procedure holds • Make sense of the overall procedure |
| Supporting Procedural Flexibility | <p>The degree to which teachers (or students) teach procedures in a way that affords students the opportunity to develop procedural flexibility, for example identifying what procedure to apply and knowing when to apply them. In supporting flexibility, the teacher or students may:</p> <ul style="list-style-type: none"> • Note multiple pathways through a procedure • Attend to applicability conditions of a procedure • Attend to key conditions of steps within a procedure • Use a procedure to conjecture about general rules • Compare multiple procedures for their affordances/limitations |
| Organization in the Presentation of the Procedure | <p>How complete, detailed, correct, and organized the teacher's (or students') presentation of content is when outlining or describing procedures, or describing the steps of a procedure used to solve problems.</p> |

Table 3

*Descriptions of the Codes in the Leveraging Connections Domain in QIPA***Leveraging Connections:**

These codes are intended to capture the connections that teachers and students forge between and among algebraic concepts, representations, and abstract ideas

| | |
|--|---|
| Connecting Across Representations | The nature of the connections teachers or students make between and across representational forms of graph, table, equation/symbolic, and context of algebraic problems, ideas, and concepts. |
| Situating the Mathematics | <p>The depth of the connections made to other aspects of the algebra curriculum, related topics, or the broader domain of mathematics, situating and motivating the current area under study within a broader context. For example:</p> <ul style="list-style-type: none"> • Making connections to prior content by linking what the class is currently doing to other topics the class has studied • Making connections to future content by articulating how what the class is currently learning sets up other topics under study • Using the architecture of the lesson to develop a mathematical through line • Making connections to other mathematical topics or the broader domain of mathematics |
| Connections Between Concrete and Abstract Ideas in Algebra | <p>The degree to which the teacher or students leverage concrete examples, representations or ideas to develop understanding of abstract concepts, formulas, notation, and definitions. They may do one or more of the following:</p> <ul style="list-style-type: none"> • Unpack the components of a formula by attending to what the symbols in the formula represent • Represent an abstract situation or idea symbolically • Clarify mathematical definitions or abstract algebraic concepts using concrete examples • Leverage abstract algebraic tools or representations to solve concrete examples • Explicitly relate an abstract concept to its analogous concrete idea • Use concrete, or pictorial examples or manipulatives to introduce or illustrate abstract ideas or concepts • Use numbers to develop generalized rules and properties |

Table 4

Descriptive Statistics for Each Code at the Segment-Level and for the Whole-Lesson Codes at the Lesson-Level

| <i>Code</i> | <i>M</i> | <i>SD</i> | <i>Min</i> | <i>Max</i> | <i>N</i> |
|--|----------|-----------|------------|------------|----------|
| Making Sense of Procedures | 1.43 | 0.75 | 1 | 5 | 261 |
| Supporting Procedural Flexibility | 1.79 | 1.00 | 1 | 5 | 261 |
| Organization in the Presentation of the Procedures | 2.89 | 0.87 | 1 | 5 | 261 |
| Connecting Across Representations | 1.69 | 1.10 | 1 | 5 | 465 |
| Situating the Mathematics | 1.18 | 0.56 | 1 | 5 | 465 |
| Connecting Between Concrete and Abstract Ideas | 1.77 | 1.13 | 1 | 5 | 465 |
| Inquiry/Exploration (Lesson-level) | 1.52 | 1.02 | 1 | 5 | 73 |
| Teacher-Led Instruction (Lesson-level) | 3.05 | 0.85 | 1 | 5 | 73 |

Table 5

Percent of Variation Attributable to Teachers Across Lessons and Across Teachers for any Given Teacher/Lesson/Segment Combination

| Code | Within-Teacher, Across Lessons | Across Teachers |
|--|-----------------------------------|-----------------|
| Making Sense of Procedures | 0.10 | 0.19 |
| Supporting Procedural Flexibility | 0.08 | 0.09 |
| Organization in the Presentation of the Procedures | 0.23 | 0.13 |
| Connecting Across Representations | 0.33 | 0.14 |
| Situating the Mathematics | 0.32 | 0.002 |
| Connecting Between Concrete and Abstract Ideas | 0.15 | 0.11 |

Table 6

Consistency of Raters for Double-Scored Lessons at the Segment Level (n = 88)

| | % Agreement | % Adjacent Agreement | % Disagreement | Kappa |
|---|-------------|----------------------|----------------|--------|
| Making Sense | 55.68 | 95.45 | 4.55 | 0.3725 |
| Supporting Flexibility | 69.32 | 93.18 | 6.82 | 0.4844 |
| Organization | 69.32 | 93.18 | 6.82 | 0.6529 |
| Overall Procedures | 65.91 | 95.45 | 4.55 | 0.5762 |
| Connecting Between Representations | 66.67 | 88.89 | 11.11 | 0.6053 |
| Situating the Mathematics | 71.60 | 96.30 | 3.70 | 0.5209 |
| Connections between Concrete and Abstract Ideas | 64.20 | 85.19 | 14.71 | 0.3629 |
| Overall Connections | 48.15 | 87.65 | 12.34 | 0.4112 |

Table 7

Consistency of Raters for Double-Scored Lessons at the Lesson-Level (n = 15)

| | % Agreement | % Disagreement | Kappa |
|---|-------------|----------------|--------|
| Making Sense | 53.33 | 0.00 | 0.3137 |
| Supporting Flexibility | 66.67 | 0.00 | 0.4828 |
| Organization | 86.67 | 0.00 | 0.7692 |
| Overall Procedures | 86.67 | 0.00 | 0.7581 |
| Connecting Between Representations | 60.00 | 0.00 | 0.5833 |
| Situating the Mathematics | 66.67 | 6.67 | 0.3919 |
| Connections between Concrete and Abstract Ideas | 73.33 | 6.67 | 0.4898 |
| Overall Connections | 38.46 | 0.00 | 0.3778 |
| Inquiry/Exploration | 76.92 | 0.00 | 0.7123 |
| Teacher-Led Instruction | 42.86 | 0.00 | 0.2899 |

Table 8

Pearson Correlations between Teacher-Level Scores on QIPA Codes and Teachers' CKT scores (n=23)

| | (1) CKT | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|----------------------------|-----------------|----------------|-----------------|------------------|--------|-----------------|--------|-----------------|--------|
| (1) CKT | 1.0000 | | | | | | | | |
| (2) Making Sense | 0.2175 | 1.0000 | | | | | | | |
| (3) Supporting Flexibility | -0.2207 | 0.3379 | 1.0000 | | | | | | |
| (4) Organization | -0.4351* | 0.1540 | 0.7638** | 1.0000 | | | | | |
| | | | * | | | | | | |
| (5) Overall Procedures | -0.3578 | 0.4950* | 0.8699** | 0.8729*** | 1.0000 | | | | |
| | | | * | | | | | | |
| (6) Representations | 0.1626 | 0.4058 | 0.2632 | 0.1175 | 0.2283 | 1.0000 | | | |
| (7) Situating | -0.1388 | 0.4151 | -0.0085 | 0.0041 | 0.0527 | -0.0675 | 1.0000 | | |
| (8) Concrete/Abstract | 0.5182* | 0.4734* | 0.3005 | 0.0547 | 0.2739 | 0.6992** | - | 1.0000 | |
| | | | | | | * | 0.1867 | | |
| (9) Overall Connections | 0.3075 | 0.5344* | 0.3235 | 0.1588 | 0.3111 | 0.8964** | 0.0082 | 0.8806** | 1.0000 |
| | | | | | | * | | * | |
| (10) Inquiry/Exploration | 0.6113* | | | | | | | | |
| | * | | | | | | | | |
| (11) Teacher-Led | -0.4465* | | | | | | | | |

*p<0.05; **p<0.01; ***p<0.001

Table 9

Pearson Correlations Between QIPA Codes and MQI Codes on Lessons with MQI Scores (n=53)

| QIPA Code | MQI Dimension | | | | | Overall MQI |
|------------------------|-----------------------|-----------------------------|-------------------------|---|-------------------------------------|----------------|
| | Errors & Imprecisions | Explicitness & Thoroughness | Richness of Mathematics | Student Participation in Meaning-Making & Reasoning | Working With Students & Mathematics | |
| Making Sense | 0.0809 | 0.0453 | 0.2110 | 0.1377 | 0.1003 | -0.0574 |
| Supporting Flexibility | 0.0272 | -0.1064 | 0.0166 | 0.2716 | 0.1915 | -0.0520 |
| Organization | 0.1115 | -0.0963 | 0.2601 | 0.1881 | 0.0346 | -0.0144 |
| Overall Procedures | 0.0638 | 0.1513 | 0.3975** | 0.3622* | 0.1834 | 0.0639 |
| Representations | 0.1716 | 0.2902* | 0.4006** | 0.2422 | 0.0027 | 0.2870* |
| Situating | 0.1462 | 0.3520* | 0.4135** | 0.3576** | 0.2395 | 0.2072 |
| Concrete/Abstract | 0.1293 | 0.2276 | 0.2007 | 0.1256 | -0.1297 | 0.0805 |
| Overall Connections | 0.2211 | 0.2686 | 0.3077* | 0.1938 | -0.0762 | 0.1161 |
| Inquiry/Exploration | -0.1757 | 0.0806 | -0.0069 | 0.0425 | -0.0017 | 0.0217 |
| Teacher-Led | 0.3138* | 0.1577 | 0.0792 | 0.0867 | -0.0948 | 0.1965 |

*p<0.05; **p<0.01; ***p<0.001

Table 10

Percentage of Lessons at Each Score Point on Format of Instruction Codes (n=108).

| | Teacher-Led Instruction | Inquiry/Exploration |
|----------|-------------------------|---------------------|
| 1 (Low) | 2 | 86 |
| 2 | 11 | 4 |
| 3 (Mid) | 54 | 6 |
| 4 | 31 | 4 |
| 5 (High) | 2 | 0 |
| Total | 100 | 100 |

Table 11

Percentage of Lessons with at Least One Segment at Each Score Level for the Teaching Procedures Domain (n=102 lessons)

| | Making Sense of Procedures | Supporting Procedural Flexibility | Organization in the Presentation of Procedures |
|-------------------------------------|-------------------------------|---|--|
| Scores Above Low (≥ 2) | 69 | 86 | 99 |
| Scores Mid or above (≥ 3) | 37 | 56 | 90 |
| Score of 4 or above (≥ 4) | 10 | 26 | 25 |
| Score of 5 | 1 | 8 | 4 |

Table 12

Percentage of Lessons with at Least One Segment at Each Score Level for the Leveraging Connections Domain (n=108 lessons)

| | Connections Across Representations | Situating the Mathematics | Connections Between Concrete and Abstract Ideas |
|-------------------------------------|---------------------------------------|------------------------------|---|
| Scores Above Low (≥ 2) | 56 | 39 | 71 |
| Scores Mid or above (≥ 3) | 46 | 12 | 56 |
| Score 4 or above (≥ 4) | 29 | 6 | 27 |
| Score 5 | 6 | 2 | 7 |

Figures

| | | | |
|---------------------------------|------------------------------------|---|---|
| Example: | Multiply second equation by 4: | Rewrite equations and solve for x: | Solve for y by plugging in $x = -3$: |
| $4x - 4y = 8$ $-8x + y = 19$ | $4x - 4y = 8$ $4(-8x + y = 19)$ | $-32x + -4y = 76$ $4x - 4y = 8$ $\hline -28x = 84$ $-28 \quad -28$ <div style="border: 1px solid black; display: inline-block; padding: 2px;">$x = -3$</div> | $4(-3) - 4y = 8$ $-12 - 4y = 8$ $+12 \quad +12$ $\hline -4y = 20$ <div style="border: 1px solid black; display: inline-block; padding: 2px;">$y = 5$</div> |

Figure 1. Systems of equations example solved by multiplying the bottom equation by four, using the elimination method to solve for x , and substituting to solve for y .

$x^2 + 6x + 9$

x

$+$

3

| | |
|-------|------|
| x^2 | $3x$ |
| $3x$ | 9 |

$(x+3)(x+3)$

Figure 2. Factoring a trinomial using the box method.

$$\begin{array}{l}
 3x - 5y = 11 \\
 x = 2y - 4 \\
 \downarrow \\
 3(2y - 4) - 5y = 11 \\
 \boxed{6y} - 12 - \boxed{5y} = 11 \\
 \downarrow \quad \swarrow \\
 y - 12 = 11 \\
 \begin{array}{r}
 y - 12 = 11 \\
 +12 \quad +12 \\
 \hline
 y = 23
 \end{array}
 \end{array}$$

$x = 2(23) - 4$
 $x = 46 - 4$
 $x = 42$

Solution: (42, 23)

1. Identify the variable that has a coefficient of 1 (i.e. y or x)
2. Replace the variable and solve
3. Replace # in the original equation

Figure 3. Worked example with accompanying steps

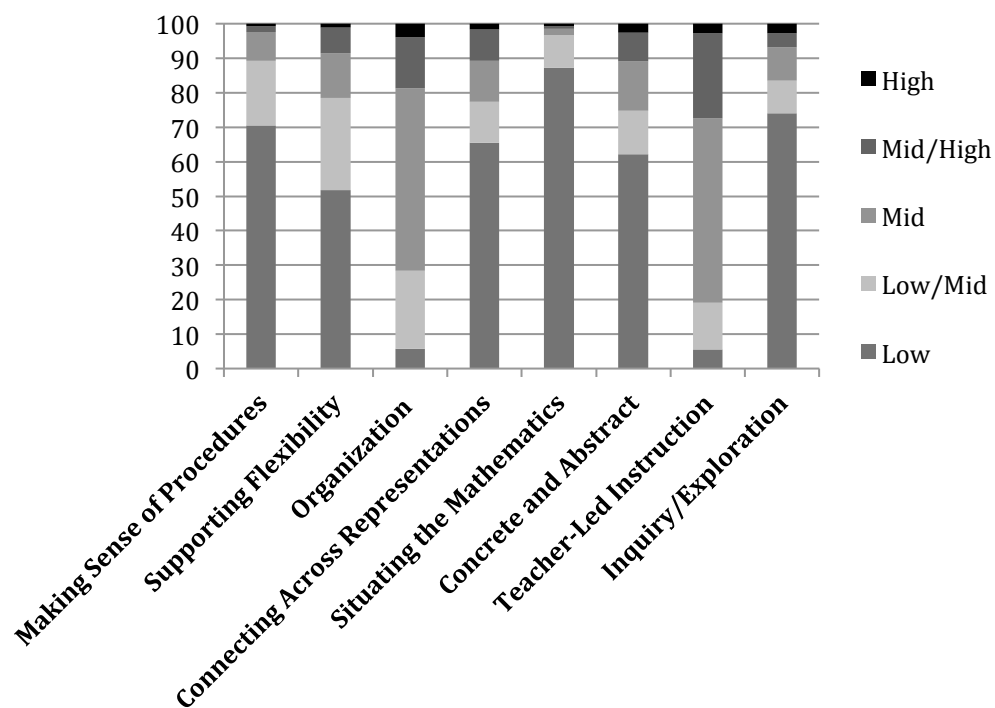


Figure 4. Percentage of scores at each score point for each code at the segment level

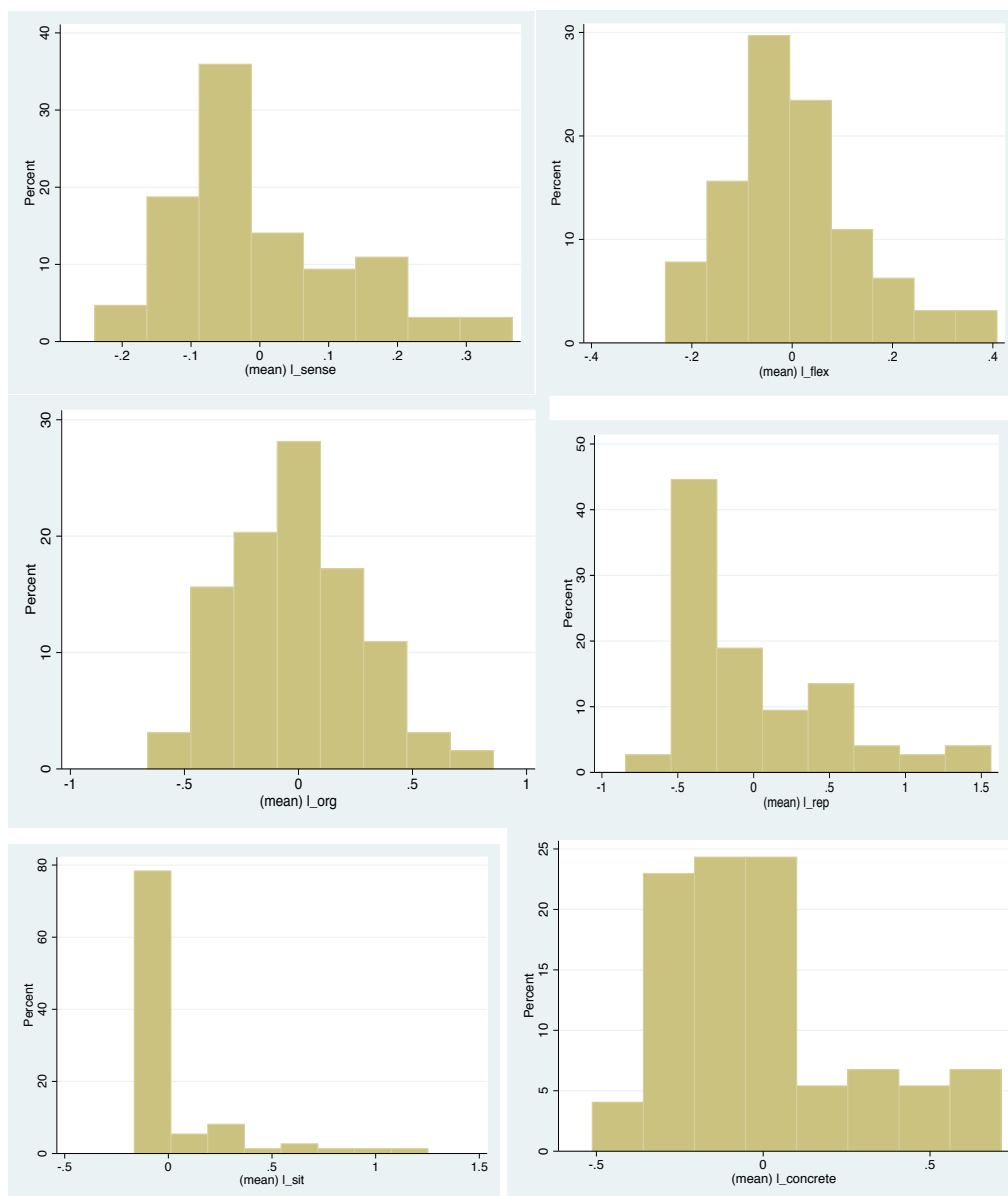


Figure 5. Distribution of shrunken averages for each code at the lesson-level (n=73) for Making Sense of Procedures and Supporting Procedural Flexibility (Row 1), Organization in the Presentation of Procedures and Connecting Across Representations (Row 2), and Situating the Mathematics and Connections Between Concrete and Abstract Ideas in Algebra (Row 3)

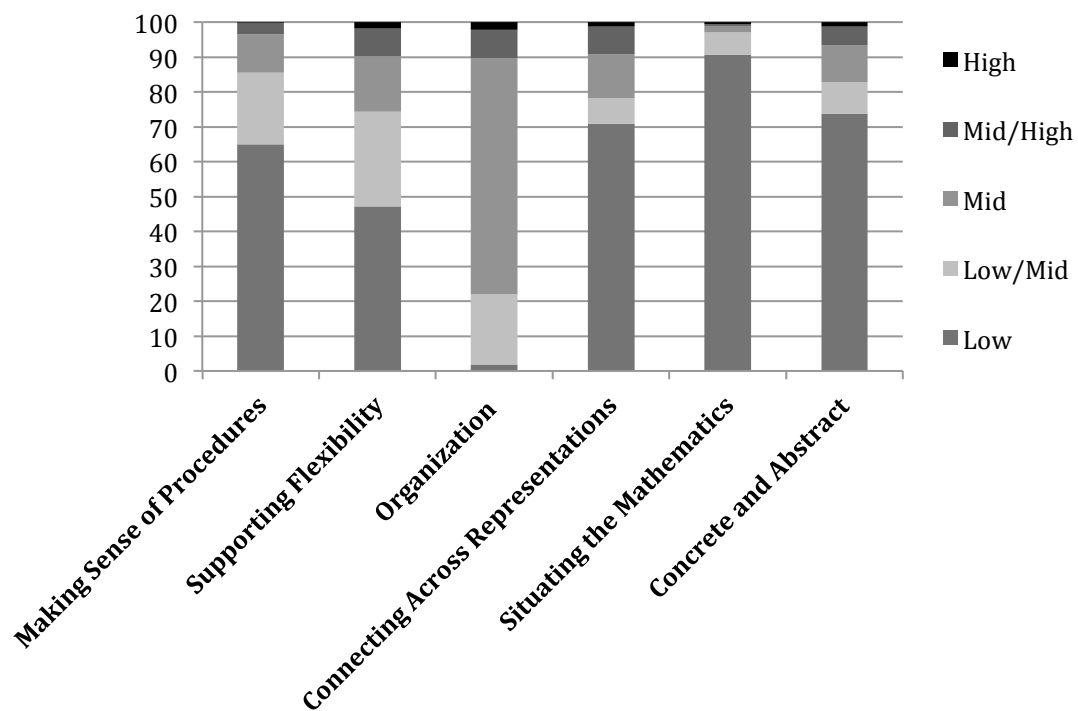
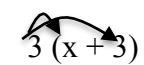


Figure 6. Percentage of segments at each score point for each code


$$3(x+3)$$

$$= 3x + 9$$

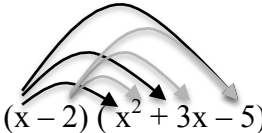

$$2x(x-2x)$$

$$= 2x^2 - 4x^2$$

$$= -2x^2$$

Figure 7. Worked examples of multiplying polynomials.

Horizontal Method:



$$\begin{aligned}
 & (x-2)(x^2+3x-5) \\
 &= x^3 + 3x^2 - 5x - 2x^2 - 6x + 10 \\
 &= x^3 + x^2 - 11x + 10
 \end{aligned}$$

Vertical Method:

$$\begin{array}{r}
 x^2 + 3x - 5 \\
 \hline
 x - 2 \\
 \hline
 -2x^2 + -6x + 10 \\
 + x^3 + 3x^2 - 5x + 0 \\
 \hline
 x^3 + x^2 - 11x - 10
 \end{array}$$

Figure 8. Horizontal and vertical methods for multiplying a binomial by a trinomial

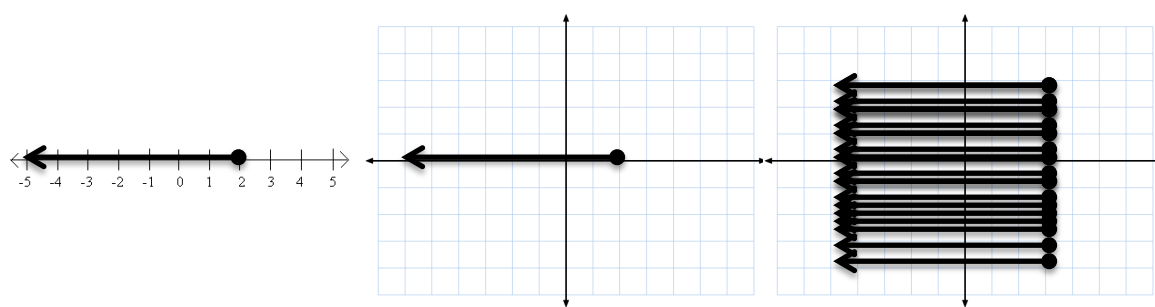


Figure 9. Graph of $x \leq 2$ on number line (Panel 1), cloned onto coordinate grid (Panel 2), and with multiple clones (Panel 3).

VITA

Erica G. Litke

| | | |
|-------------|---|-------------------|
| 1993 – 1997 | Oberlin College Oberlin, OH | B.A. May 1997 |
| 1997 – 2000 | Program Director, National Dance Institute New York, NY | |
| 2000 – 2001 | Harvard Graduate School of Education Cambridge, MA | Ed.M. May 2001 |
| 2001 – 2002 | Director of Education, Phipps W. Farms Beacon Bronx, NY | |
| 2002 – 2008 | Teacher, East Side Community High School New York, NY | |
| 2008 – 2009 | Achievement Coach, Urban Assembly New York, NY | |
| 2009 – 2015 | Doctor of Education Candidate Harvard University Graduate School of Education Cambridge, MA | |
| 2009 – 2015 | Research Assistant Harvard University Graduate School of Education Cambridge, MA | |
| 2010 – 2015 | Teaching Fellow Harvard University Graduate School of Education Cambridge, MA | |